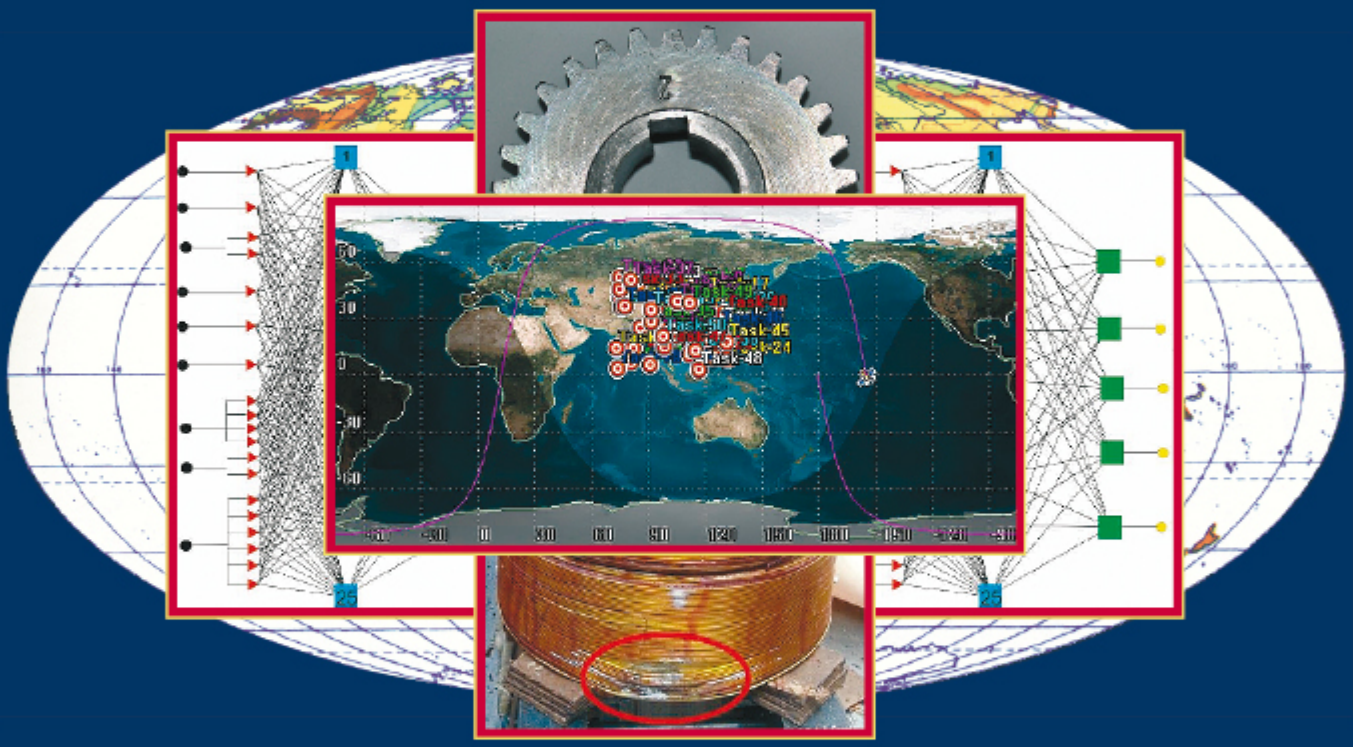
Vol. 19. No 3, 2017

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

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY



Polskie Naukowo Techniczne Towarzystwo Eksploatacyjne
Warszawa

Polish Maintenance Society
Warsaw

Co-financed by the Minister of Science and Higher Education

The Journal is indexed and abstracted in the Journal Citation Reports (JCR Science Edition), Scopus, Science Citation Index Expanded (SciSearch®) and Index Copernicus International.

The Quarterly appears on the list of journals credited with a high impact factor by the Polish Ministry of Science and Higher Education and is indexed in the Polish Technical Journal Contents database – BAZTECH and the database of the Digital Library Federation.

All the scientific articles have received two positive reviews from independent reviewers.

Our 2015 Impact Factor is 1.248

Editorial staff:	Dariusz Mazurkiewicz, PhD, DSc (Eng), Associate Professor (Editor-in-Chief, Secretary of the Scientific 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

PIETRUCHA-URBANIK K, STUDZIŃSKI A. **Case study of failure simulation of pipelines conducted in chosen water supply system**. Eksploatacja i Niezawodnosc – Maintenan ce and Reliability 2017; 19 (3): 317–323, http://dx.doi.org/10.17531/ein.2017.3.1.

The main goal of this work is to simulate the failure of water pipe network, using the hydraulic model of the network created through Epanet 2 program. The model includes the cooperation of the second stage pumping station with the expansion tanks located in the network. Based on these parameters, the simulation operation of water supply network was performed, as well as failure simulation on the basis of closing some sections of water pipe network. Failure analysis allowed to perform characteristics of the water supply system including pressure changes that occur in the network during failure simulation.

PIETRUCHA-URBANIK K, STUDZIŃSKI A. **Symulacje awarii rurociągów przeprowadzone na wybranej sieci wodociągowej – analiza przypadku**. Eksploatacja i Niezawodnosc – Maintenan ce and Reliability 2017; 19 (3): 317–323, http://dx.doi.org/10.17531/ein.2017.3.1.

Głównym celem pracy jest przeprowadzenie symulacji awarii przewodów przewodów wodociągowych, za pomocą stworzonego modelu hydraulicznego sieci z wykorzystaniem programu Epanet 2. W modelu zawarta jest współpraca pompowni drugiego stopnia ze zbiornikami wyrównawczymi znajdującymi się na sieci. Bazując na tych parametrach przeprowadzono symulację pracy sieci wodociągowej, a dzięki wyłączeniem odcinków przewodów wodociągowych symulację awarii. Analizowanie awarii pozwoliło na scharakteryzowanie dostawy wody do odbiorców w tym zmian, jakie występują na sieci podczas symulowania awarii.

GODZIMIRSKI J, ROŚKOWICZ M, KOMOREK A. **Investigation of ballistic resistance of adhesive bonded multi-layer structures**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 324–330, http://dx.doi.org/10.17531/ein.2017.3.2.

The aim of the study was to evaluate the possibility of using multilayer structures for light armour, in which the elements of the ceramic type SiC and Al_2O_3 were bonding by adhesive joins with antiballistic polyethylene and aramid fabrics. Ballistic resistance shells was determined using for this purpose the test stand constructed on the basis of helium gun and high-speed camera. It has been shown that the laminated structure composed of thin metal layers and aramid fabrics as well as polymer composites made of aramid fabric has lower ballistic resistance than loose fabric packs. It was also demonstrated the functionality of the use of the ceramic component as a separate package, which the ceramic plates are adhesive bonded between the two layers of sheet metal. There is also evidence that fabrics composed of thin layers of material poorly connected with each other, should not be adhesively bonded to the ceramic. It proposed the preparation of specimens, which best reconstruct the load of ceramic plates adhesive bonded to fabric, which are made of lightweight bulletproof vests and ballistic panels.

GODZIMIRSKI J, ROŚKOWICZ M, KOMOREK A. **Badania odporności na przebicie klejonych struktur wielowarstwowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 324–330, http://dx.doi.org/10.17531/ein.2017.3.2.

Celem badań była ocena możliwości zastosowania opracowanych struktur wielowarstwowych do wytwarzania lekkich pancerzy, mogących znaleźć zastosowanie jako kuloodporne osłony balistyczne śmigłowców bojowych i innego lekkiego sprzętu wojskowego. Istotnym celem badań było również sprawdzenie możliwości łączenia metodą klejenia ceramiki typu SiC i Al_2O_3 z antybalistycznymi tkaninami polietylenowymi i aramidowymi. Odporność na przebicie była badana z wykorzystaniem stanowiska z działem helowym i ultraszybką kamerą. W wyniku przeprowadzonych badań stwierdzono, iż pakiety luźnych tkanin aramidowych charakteryzuje większa odporność na przebicie w porównaniu z kompozytami polimerowymi wytworzonymi z takich tkanin oraz kompozytami warstwowymi złożonymi z cienkich warstw metalowych i tkanin aramidowych. Ponadto zauważono, że warstwy ceramiczne znacznie zwiększają odporność na przebicie osłon balistycznych i ich stosowanie w takich osłonach wydaje się niezbędne oraz, że wklejenie płytek ceramicznych pomiędzy dwie cienkie blachy ze stopu aluminium i nie łączenie ich bezpośrednio z tkaninami aramidowymi zapewnia w przypadku uderzenia pociskiem zniszczenie małej powierzchni warstwy ceramicznej, a więc w małym stopniu zmniejsza właściwości ochronne osłony.

STĘPIEŃ Z. **A study of factors influencing the formation of harmful deposits in the diesel engine injectors**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 331–337, http://dx.doi.org/10.17531/ein.2017.3.3.

The paper describes various types of deposits of injectors used in CI engines with direct and indirect fuel injection. Various factors have been determined which affect the formation and size of deposits. These factors have also been classified and evaluated in terms of their impact on the formation of deposits precursors and further changes of deposits during the engine life. The results of deposits formation tests have been presented. The external and internal deposits in the High Pressure Common Rail fuel injection systems have been photographed and described. The impact of formed deposits on the diagnostic parameters of injectors has been discussed broadly as it determines the range of injectors' dysfunction and their fit for further use.

STĘPIEŃ Z. **Badanie czynników wpływających na tworzenie szkodliwych osadów wtryskiwaczy paliwa silników z zapłonem samoczynnym**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 331–337, http://dx.doi.org/10.17531/ein.2017.3.3.

W artykule opisano różne typy osadów wtryskiwaczy silników ZS z pośrednim i bezpośrednim wtryskiem paliwa. Określono różnorodne czynniki mające wpływ na tworzenie i zmiany wielkości osadów. Dokonano też klasyfikacji i oceny znaczenia przedmiotowych czynników na powstawanie prekursorów osadów oraz dalsze zmiany osadów w czasie eksploatacji silnika. W zakresie opisu przeprowadzonych badań przedstawiono wyniki procesów tworzenia osadów w warunkach testów silnikowych. Utworzone w ten sposób osady zewnętrzne i wewnętrzne wtryskiwaczy układu wysokociśnieniowego wtrysku paliwa typu common rail były następnie udokumentowane fotograficznie, oraz opisane. Szeroko omówiono też wyniki oceny wpływu wytworzonych osadów wtryskiwaczy na zmiany ich parametrów diagnostycznych, określających zakres dysfunkcji wtryskiwaczy i decydujących o możliwości ich dalszego stosowania.

HAWRYLUK M, MRZYGŁÓD B. **A durability analysis of forging tools for different operating conditions with application of a decision support system based on artificial neural networks (ANN)**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 338–348, http://dx.doi.org/10.17531/ein.2017.3.4.

The paper presents the results of research concerning the percentage participation of destructive mechanisms for two typical variants of exploitation of forging tools: lubricated and cooled, and without lubrication. Discussed results come from the developed by the authors the decision support system (SEPEK-2) based on artificial neural network. The knowledge about the durability of forging tools needed for learning artificial neural network was included in the training data set, from comprehensive studies, carried out in industrial conditions. Set of training data set included 450 records of knowledge. The paper presents the process of acquiring knowledge, adopted neural network architecture and parameters developed network. Carried out a global analysis of the results generated by the developed system for the durability of forging tools treated as the maximum number of produced forgings to their destruction (from 0 to 25,000 items), showed that for the lubricated and cooling tools the dominant mechanism is thermo-mechanical fatigue, and do not abrasive wear, which actually dominates in the process of forging tools for uncooled and unlubricated tools. It should be emphasized that the overwhelming majority of studies in this area is attributed that to abrasive wear is dominant, and as shown by the results of research and analysis for the selected representative forging processes, with the use of decision support system based on ANN, the fatigue a thermo-mechanical is dominant in these processes. However, due to the easy measurability and commonly used models wear, based on the model of Archard, it is abrasive wear assigned the largest participation. In fact, for the tool lubricated and cooled tools a thermo-mechanical fatigue intensifies this

HAWRYLUK M, MRZYGŁÓD B. **Analiza trwałości narzędzi kuźniczych dla różnych warunków eksploatacji z wykorzystaniem systemu wspomagania decyzji opartego o sztuczne sieci neuronowe**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 338–348, http://dx.doi.org/10.17531/ein.2017.3.4.

W pracy przedstawiono wyniki badań, dotyczące, procentowego udziału mechanizmów niszczących dla dwóch typowych wariantów eksploatacji narzędzi kuźniczych: smarowanych i chłodzonych oraz bez smarowania. Prezentowane wyniki pochodzą z opracowanego przez autorów systemu wspomagania decyzji (SEPEK-2) działającego w oparciu o sztuczną sieć neuronową. Wiedza o analizowanym zagadnieniu trwałości narzędzi kuźniczych, potrzebna do procesu uczenia sztucznej sieci neuronowej zawarta była w zestawie danych uczących, pochodzących z kompleksowych badań, zrealizowanych w warunkach przemysłowych. Zestaw danych uczących obejmował zbiór 450 rekordów wiedzy. W pracy przestawiono proces pozyskiwania wiedzy, przyjętą architekturę sieci neuronowej oraz parametry opracowanej sieci. Przeprowadzona globalna analiza wyników generowanych przez opracowany system, dla trwałości traktowanej jako zwiększająca się liczba odkuwek (od 0 do 25000 sztuk), wykazała że dla narzędzi smarowanych i chłodzonych dominującym mechanizmem jest zmęczenie cieplno-mechaniczne, a nie zużycie ścierne, które rzeczywiście dominuje w procesach kucia dla narzędzi niechłodzonych i niesmarowanych. Należy podkreślić, że zdecydowana większość opracowań z tego obszaru przypisuje, że to zużycie ścierne jest dominujące, a jak wykazały wyniki badań i analiz dla wybranych reprezentatywnych procesów kucia, przy wykorzystaniu systemu wspomagania decyzji opartego o SNN, to zmęczenie cieplno-mechaniczne jest dominujące w tych procesach. Jednakże ze względu na

effect attributed to abrasive wear. While the generally accepted view is correct, in the case of tools unlubricated, as confirmed by the analysis using ANN.

łatwą mierzalność oraz popularnie stosowane modele zużycia ściernego, bazujące na modelu Archarda, to właśnie zużyciu ściernemu przypisuję się największy udział, choć w rzeczywistości dla narzędzi smarowanych i chłodzonych zmęczenie cieplno-mechaniczne wzmaga ów efekt przypisywany zużyciu ściernemu. Natomiast ogólnie przyjęty pogląd jest słuszny, w przypadku narzędzi niesmarowanych. co potwierdziły także analizy przy wykorzystaniu SNN.

WANG Y, BI L, WANG S, LIN S, XIANG W. **The application of dynamic bayesian network to reliability assessment of emu traction system**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 349–357, http://dx.doi.org/10.17531/ein.2017.3.5.

The article introduces a novel application of a Dynamic Bayesian Network (DBN) in the reliability assessment with regard to the traction system of Electric Multiple Units (EMU), which focus on modeling approach to DBN construction. As a result of high complexity and growing interdependencies, it is increasingly vulnerable to the failure of components. Although many studies on the use of BN for estimating the system reliability have been conducted, there is a lack of effective modeling power regarding current tools in depicting both functional and temporal dependencies between components. In this paper, a new modeling approach to DBN generation is submitted, which can be applied to the system made up of certain components and different types of flows propagating through them. The Component-based CPT (Conditional Probability Table) and Time-dependent CPT are used to describe functional dependencies and temporal dependencies respectively. As the complexity of the system cannot be modeled in a tractable way as a DBN, a Breadth-First-Search (BFS) algorithm is introduced for the construction of the DBN model in an automated manner. With the application of the proposed DBN-based approach, the reliability of the traction system can be evaluated at any given time, which is of great significance to determine the plan of maintenance in an effort to ensure the system safety.

WANG Y, BI L, WANG S, LIN S, XIANG W. **Zastosowanie dynamicznych sieci bayesowskich do oceny niezawodności elektrycznego systemu trakcyjnego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 349–357, http://dx.doi.org/10.17531/ein.2017.3.5.

W artykule omówiono nowatorskie zastosowanie dynamicznej sieci bayesowskiej (DBN) do oceny niezawodności elektrycznego systemu trakcyjnego ze szczególnym uwzględnieniem metod modelowania DBN. W związku z rosnącą złożonością elektrycznych systemów trakcyjnych oraz wynikającą z niej coraz większą ilością współzależności między komponentami, systemy te narażone są coraz częściej na awarie części składowych. Chociaż istnieje wiele badań dotyczących oceny niezawodności systemów trakcyjnych, stosowane obecnie narzędzia nie mają odpowiedniej mocy modelowania koniecznej do opisu zależności funkcjonalnych i czasowych pomiędzy częściami składowymi. W niniejszej pracy zaproponowano nową metodę modelowania generowania DBN, którą można stosować w odniesieniu do systemów składających się z pewnych określonych komponentów oraz różnych typów rozchodzących się przez nie przepływów. Zależności funkcjonalne i czasowe opisano, odpowiednio, za pomocą tablicy komponentowych prawdopodobieństw warunkowych (Component-based Conditional Probability Table, CPT) oraz tablicy czasowo-zależnych prawdopodobieństw warunkowych. Ponieważ złożoność systemu nie pozwala na zamodelowanie go w prosty sposób jako DBN, do automatycznej budowy modelu DBN wykorzystano algorytm przeszukiwania wszerz (Breadth-First-Search). Oceny niezawodności systemu trakcyjnego z wykorzystaniem proponowanej metody opartej na DBN można dokonywać w dowolnym czasie, co ma ogromne znaczenie przy planowaniu konserwacji w celu zapewnienia bezpieczeństwa systemu.

RAPOSO H, FARINHA JT, FERREIRA L, GALAR D. **An integrated econometric model for bus replacement and determination of reserve fleet size based on predictive maintenance**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 358–368, http://dx.doi.org/10.17531/ein.2017.3.6.

Maintenance policies influence equipment availability and, thus, they affect a company's capacity for productivity and competitiveness. It is important to optimize the Life Cycle Cost (LCC) of assets, in this case, passenger bus fleets. The paper presents a predictive condition monitoring maintenance approach based on engine oil analysis, to assess the potential impact of this variable on the availability of buses. The approach has implications on maintenance costs during the life of a bus and, consequently, on the determination of the best time for bus replacement. The paper provides an overview of economic replacement models through a global model, with an emphasis on availability and its dependence on maintenance and maintenance costs. These factors help to determine the size of the reserve fleet and guarantee availability.

RAPOSO H, FARINHA JT, FERREIRA L, GALAR D. **Zintegrowany ekonometryczny model do modelowania wymiany taboru autobusowego oraz określania wielkości floty rezerwowej w oparciu o konserwację predykcyjną**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 358–368, http://dx.doi.org/10.17531/ein.2017.3.6.

Polityka konserwacji wpływa na gotowość sprzętu, a tym samym na wydajność i konkurencyjność przedsiębiorstwa. Ważne jest optymalizowanie kosztów cyklu życia (LCC) aktywów, w tym przypadku taboru autobusowego. W artykule przedstawiono metodę utrzymania ruchu polegającą na predykcyjnym monitorowaniu stanu w oparciu o analizę oleju silnikowego w celu oceny potencjalnego wpływu tej zmiennej na gotowość autobusów. Podejście to ma praktyczne konsekwencje jeśli chodzi o koszty utrzymania w trakcie eksploatacji autobusu, a także pozwala na ustalenie najlepszego czasu na wymianę pojazdów taboru. W pracy przedstawiono przegląd ekonomicznych modeli wymiany oraz opracowano model globalny integrujący te modele, ze szczególnym uwzględnieniem gotowości oraz jej zależności od konserwacji oraz kosztów utrzymania ruchu. Czynniki te pomagają określić wielkość floty rezerwowej i zapewnić gotowość taboru.

SURBLYS V, ŽURAULIS V, SOKOLOVSKIJ E. **Estimation of road roughness from data of on-vehicle mounted sensors**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 369–374, http://dx.doi.org/10.17531/ein.2017.3.7.

In this work, methods of estimation of road roughness by processing data of laser sensors of distance measurement are investigated. Depending on the layout of the sensors and the processing of the data, the Japanese, symmetrical and asymmetrical roughness estimation methods are elaborated in this work. The experimental research that was carried out by "Japanese method" on two known profile road surfaces allowed evaluating the influence of the vehicle's suspension and the body pitch rate oscillation on the calculation of the roughness. The validation of the results showed that the compensation of the suspension travel does not warrant the sufficient accuracy of the estimation of the road roughness because the pivoting of the bodywork has a greater influence on the processing of the signals recorded by the sensors attached to it.

SURBLYS V, ŽURAULIS V, SOKOLOVSKIJ E. **Ocena nierówności drogowych zgodnie z danymi czujników zainstalowanych na pojeździe**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 369–374, http://dx.doi.org/10.17531/ein.2017.3.7.

Artykuł przedstawia i ocenia metody określania nierówności nawierzchni drogi na podstawie danych z laserowych czujników pomiaru odległości. W artykule omówiono opracowane w Japonii symetryczne i asymetryczne metody określania nierówności nawierzchni, które różnią się sposobem rozmieszczenia czujników i przetwarzania danych. Badania eksperymentalne przeprowadzone na podstawie „metody japońskiej" na dwóch zadanych profilach nawierzchni pozwoliły na ocenę wpływu zawieszenia samochodu i zmian przechyłu nadwozia względem osi podłużnej na obliczanie nierówności nawierzchni. Walidacja wyników pokazała, że kompensacja odchylenia zawieszenia nie zapewnia wystarczającej dokładności pomiaru nierówności drogi, ponieważ przechyły nadwozia mają większy wpływ na przetwarzanie sygnałów rejestrowanych przez zamontowane na nim czujniki.

PAŚ J, ROSIŃSKI A. **Selected issues regarding the reliability-operational assessment of electronic transport systems with regard to electromagnetic interference**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 375–381, http://dx.doi.org/10.17531/ein.2017.3.8.

The article presents issues associated with the reliability-operational analysis of electronic transport systems, which are operated in a defined environment. Intended or unintended (static or mobile) electromagnetic interference, which are present over a broad transport area, can cause interference of its functioning. That is why correct functioning of electronic transport systems in a distorted electromagnetic environment is so important. The articles measurement results of low frequency radiation basic characteristics for power supplies, which were carried out for two frequency ranges: (0÷400) Hz and (400÷6500) Hz. The impact

PAŚ J, ROSIŃSKI A. **Wybrane zagadnienia oceny niezawodnościowo-eksploatacyjnej transportowych systemów elektronicznych z uwzględnieniem zakłóceń elektromagnetycznych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 375–381, http://dx.doi.org/10.17531/ein.2017.3.8.

W artykule przedstawiono zagadnienia związane z analizą niezawodnościowo-eksploatacyjną transportowych systemów elektronicznych, które są eksploatowane w określonym środowisku. Występujące na rozległym obszarze transportowym zaburzenia elektromagnetyczne zamierzone lub niezamierzone (stacjonarne lub ruchome) mogą być przyczyną zakłócenia ich funkcjonowania. Dlatego tak istotne jest prawidłowe funkcjonowanie transportowych systemów elektronicznych w zniekształconym środowisku elektromagnetycznym. W artykule przedstawiono wyniki pomiarów podstawowych charakterystyk promieniowania niskiej częstotliwości dla zasilaczy, które zostały przeprowadzone dla

of the load's voltage change on the electric field E [V/m] produced by the power supply and the impact of power supply load's current changes on the induction of the magnetic field B [nT] were presented. Next, a graph of relations in an electronic transport system, with regard to level of interference in electric and magnetic fields, was developed. This allowed to determine the relations, which enable to define the probability values of a system being in the distinguished states. The methodology of the reliability-operational analysis of electronic transport systems with regard to electromagnetic interference presented in this article, may be used during designing of electronic systems used in transport.

dwóch zakresów częstotliwości: (0÷400) Hz i (400÷6500) Hz. Zaprezentowano wpływ zmiany napięcia obciążenia na pole elektryczne E [V/m] wytwarzane przez zasilacz, oraz wpływ zmiany prądu obciążenia zasilacza na indukcję pola magnetycznego B [nT]. Następnie opracowano graf relacji w transportowym systemie elektronicznym z uwzględnieniem poziomów zakłóceń pola elektrycznego i magnetycznego. Umożliwiło to wyznaczenie zależności pozwalających na określenie wartości prawdopodobieństw przebywania systemu w wyróżnionych stanach. Zaprezentowana w artykule metodyka analizy niezawodnościowo-eksploatacyjnej transportowych systemów elektronicznych z uwzględnieniem zakłóceń elektromagnetycznych może być użyta podczas projektowania systemów elektronicznych stosowanych w transporcie.

JACYNA-GOŁDA I, LEWCZUK K. **The method of estimating dependability of supply chain elements on the base of technical and organizational redundancy of process**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 382–392, http://dx.doi.org/10.17531/ein.2017.3.9.

The quality of supply chain operation depends on quality of its particular elements, including warehouses. The paper presents an attempt to describe the quality of warehouse operation in terms of dependability. Authors discussed issues related to assessing warehouse operation, quality problems and solutions to increase the quality of work. The technical and organizational redundancy was proposed as a primary factor increasing dependability of warehous operation in supply chain and thereby improving the quality of services. Authors discussed dependability of supply chain and warehouses and have proposed an approach to determination of dependability of warehouse facility based on technological and organisational redundancy related to material flow pile-ups. The approach was founded on OTIFEF index as a base for dependability estimation. Construction of that index basing on probabilities of correct realization of different aspects of logistics service was proposed. An important element of the approach presented in the paper is proposal of technical and organisational indicators defining different aspects of redundancy in aspect of dependability. The example of redundancy assessment in function of technical and organisational methods of increasing warehouse efficiency has been provided.

JACYNA-GOŁDA I, LEWCZUK K. **Metoda szacowania niezawodności elementów łańcucha dostaw na podstawie charakterystyk nadmiarowości technicznej i organizacyjnej procesu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 382–392, http://dx.doi.org/10.17531/ein.2017.3.9.

Jakość pracy łańcucha dostaw jest wynikiem jakości pracy jego elementów, w tym magazynów. W artykule przedstawiono próbę ujęcia zagadnień jakościowych pracy magazynu w kategoriach niezawodnościowych. Omówiono zagadnienia oceny pracy magazynów, źródła problemów jakościowych i stosowane rozwiązania zwiększające jakość pracy. Wskazano nadmiarowość techniczną i organizacyjną jako podstawowy środek zwiększania niezawodności realizacji zadań przez magazyny w łańcuchu dostaw i tym samym poprawę jakości świadczonych usług. Autorzy omówili zagadnienia niezawodności łańcucha dostaw i magazynów jako ich podstawowych elementów i zaproponowali podejście do określania niezawodności magazynu oparte o nadmiarowość technologiczną i organizacyjną ustalaną w oparciu o przewidywane spiętrzenia przepływu materiałów. Do tego celu wykorzystano miernik OTIFEF jako podstawę szacowania niezawodności. Zaproponowano konstrukcję tego miernika w oparciu o prawdopodobieństwa poprawnej realizacji różnych aspektów usług logistycznych. Ważnym elementem podejścia proponowanego w artykule jest propozycja technicznych i organizacyjnych wskaźników określających różne aspekty nadmiarowości w funkcji niezawodności magazynu. Przedstawiono przykład szacowania nadmiarowości z wykorzystaniem technicznych i organizacyjnych metod zwiększania efektywności.

GHANOONI-BAGHA M, SHAYANFAR MA, REZA-ZADEH O, ZABIHI-SAMANI M. **The effect of materials on the reliability of reinforced concrete beams in normal and intense corrosions**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 393–402, http://dx.doi.org/10.17531/ein.2017.3.10.

Concrete structures are exposed to a variety of damages during their lifetime each of which could contribute to reducing their service life and load bearing capacity. Since most of parameters have special role in estimating capacity of members which are not certain, probabilistic evaluating the performance of concrete structures could bring more realistic perception about analysis and design of these structures. One of the most frequent probable damages is corrosion. The main focus of this study is placed on reliability assessment of flexural behavior of a reinforced concrete beam experienced pitting corrosion via Monte Carlo simulation. In addition, the effects of time to corrosion initiation, steel rebar diameter, yielding stress of rebars, strength class of cement, aggregate type and compressive strength of concrete, are included both in intense and normal pitting corrosion. The results clearly illustrate that occurrence of intense corrosion in concrete with low compressive strength, which used of higher strength class of cement and crushed stone aggregate, and less initial time for corrosion will lead to considerable reduction in service life even in some cases nearly half.

GHANOONI-BAGHA M, SHAYANFAR MA, REZA-ZADEH O, ZABIHI-SAMANI M. **Wpływ stosowanych materiałów na niezawodność belek żelbetowych w warunkach normalnej i silnej korozji**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 393–402, http://dx.doi.org/10.17531/ein.2017.3.10.

W trakcie cyklu życia, konstrukcje betonowe są narażone na wiele uszkodzeń, z których każde może przyczyniać się do skrócenia ich żywotności i nośności. Ponieważ większość parametrów odgrywających szczególną rolę w szacowaniu nośności elementów cechuje niepewność, ocena probabilistyczna charakterystyk struktur betonowych może dawać bardziej realistyczny obraz analizy i projektowania tych struktur. Jednym z najczęściej występujących uszkodzeń struktur żelbetowych jest korozja. Głównym celem niniejszego badania była ocena niezawodności zachowania zginanej belki żelbetowej doświadczalnie poddanej korozji wżerowej poprzez symulację Monte Carlo. Ponadto, badano oddziaływanie czasu inkubacji korozji, średnicy stalowych prętów zbrojeniowych, granicy plastyczności tych prętów, klasy wytrzymałości cementu, rodzaju kruszywa i wytrzymałości na ściskanie betonu zarówno w warunkach silnej jak i normalnej korozji wżerowej. Wyniki jasno pokazują, że wystąpienie silnej korozji w betonie o małej wytrzymałości na ściskanie, do produkcji którego wykorzystano cement i kruszywo kamienne o wyższej klasie wytrzymałości, oraz krótszy czas inkubacji korozji prowadzą do znacznego skrócenia żywotności belek, w niektórych przypadkach nawet prawie o połowę.

JODEJKO-PIETRUCZUK A, WERBIŃSKA-WOJCIECHOWSKA S. **Development and sensitivity analysis of a technical object inspection model based on the delay-time concept use**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 403–412, http://dx.doi.org/10.17531/ein.2017.3.11.

In the presented paper, authors focus on the development of mathematical delay-time model for single-unit technical systems (technical objects) liable to costly failure. Failure is taken here to mean a breakdown or catastrophic event, after which the system is unusable until replacement. Implemented maintenance policy is the Block-Inspection Policy that assumes performing inspection actions at regular time intervals of T. In the perfect inspection case the availability and cost models are developed. This gives the possibility for analytical optimization of time between maintenance actions performance T for the infinite operational time of the system. Later, there is examined the compatibility of the developed analytical model with simulation results. The main target is to investigate what is the influence of the given model basic time components on the system availability ratio level and the system long-run expected maintenance costs. The analysis is conducted in the two main steps. The first one regards to analysis of expected number of events (failures, preventive replacements and inspection actions) in a single renewal cycle for the chosen range of time parameters: T and delay time h. In the next step, the availability ratio and long-run maintenance costs dependency on the chosen model's time parameters is under consideration. At the end, the directions for further research work are defined.

JODEJKO-PIETRUCZUK A, WERBIŃSKA-WOJCIECHOWSKA S. **Opracowanie i analiza wrażliwości modelu kontroli stanu obiektu technicznego z wykorzystaniem koncepcji opóźnień czasowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 403–412, http://dx.doi.org/10.17531/ein.2017.3.11.

W artykule autorzy skupili się na opracowaniu matematycznego modelu utrzymania obiektów technicznych podlegających kosztownym uszkodzeniom z uwzględnieniem koncepcji opóźnień czasowych. Uszkodzenie w danym przypadku oznacza awarię lub zdarzenie katastrofalne, po którym obiekt jest niezdatny do użytku do momentu wymiany. Wykorzystano politykę blokowej kontroli stanu obiektu, która zakłada, że operacje diagnozy jego stanu są przeprowadzane w regularnych odstępach co T jednostek czasu. Rozpatrzono model kosztowy oraz model gotowości dla przypadku perfekcyjnej diagnozy stanu obiektu. Pozwoliło to na przeprowadzenie analitycznej optymalizacji okresu T między kolejnymi diagnozami stanu obiektu dla nieskończonego horyzontu czasowego. Następnie, zbadano zgodność opracowanego modelu analitycznego z wynikami uzyskanymi w drodze symulacji. Głównym celem było zbadanie wpływu podstawowych parametrów czasowych opracowanego modelu na poziom współczynnika gotowości oraz oczekiwanych kosztów utrzymania badanego obiektu. Analiza została przeprowadzona w dwóch etapach. Pierwszy obejmuje analizę oczekiwanej liczby zdarzeń (uszkodzeń, wymian profilaktycznych oraz operacji kontroli stanu obiektu) dla wybranych zakresów parametrów czasowych: T i opóźnienia czasowego h. W kolejnym kroku zbadano zależność wskaźnika gotowości i oczekiwanych kosztów utrzymania obiektu od wybranych parametrów czasowych modelu. Pracę kończy wskazanie kierunków dalszych prac badawczych.

PAVLOVIĆ P, MAKAJIĆ-NIKOLIĆ D, VUJOŠEVIĆ M. **A new approach for determining the most important system components and the budget-constrained system reliability improvement**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 413–419, http://dx.doi.org/10.17531/ein.2017.3.12.

Importance measures are used for indexing system components due to their impact on the system's overall reliability. In order to identify the specific number of the most critical components, first-ranked components are singled out as the most important ones. However, importance measures consider only the influence of individual components and they are not applicable to combinations or groups of components. This common feature of importance measures is referred to in literature as one of still open issues. This paper proposes a new approach for determining the most important system components, where a whole set of components are determined simultaneously taking into account their interdependence. In systems with a large number of interdependent components, the number of the most important components which should be prevented is often limited due to the available budget. Using pre-known minimal cut sets, a mathematical model based on the Budgeted Maximum Coverage Problem is proposed. By its optimization, the simultaneous determination of all of the most important components whose total expenses do not exceed the limited overall budget is achieved. The new approach was tested by a series of experiments conducted over a set of test examples. The results of the experiments were compared with the results obtained using two commonly used cost importance measures.

PAVLOVIĆ P, MAKAJIĆ-NIKOLIĆ D, VUJOŠEVIĆ M. **Nowe podejście do wyznaczania najważniejszych elementów systemu oraz poprawy niezawodności systemu w warunkach ograniczonego budżetu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 413–419, http://dx.doi.org/10.17531/ein.2017.3.12.

W złożonych systemach, w których koszty poprawy niezawodności poszczególnych elementów są znane, często ogranicza się budżet przeznaczony na podnoszenie ogólnej niezawodności systemu. W takich przypadkach konieczna jest maksymalizacja niezawodności systemu przy jednoczesnym utrzymaniu kosztów na poziomie minimum. Powszechnie znane metody rozwiązywania powyższego problemu opierają się na wyznaczaniu ważności kosztów, co wymaga określenia rang elementów składowych systemu, a w dalszej kolejności wyodrębnienia pewnej liczby najważniejszych elementów pierwszorzędnej rangi. W niniejszej pracy zaproponowano nowe podejście do określania najważniejszych komponentów systemu w oparciu o problem maksymalnego pokrycia w granicach budżetu (budgeted maximum coverage problem); podejście wdrażano z wykorzystaniem wcześniej znanych minimalnych przekrojów niezdatności. Optymalizacja proponowanego modelu matematycznego, pozwoliła na jednoczesne wyznaczenie wszystkich najważniejszych elementów, dla których łączne wydatki na utrzymanie ruchu nie przekraczały całkowitego ograniczonego budżetu. Nowe podejście zostało przebadane w serii eksperymentów przeprowadzonych na zbiorze przykładów testowych, za które posłużyły wzorcowe drzewa błędów. Wyniki badań porównano z wynikami uzyskanymi za pomocą dwóch miar ważności kosztów – miary ważności opartej na kosztach oraz miary ważności opartej na opłacalności. W większości przypadków, proponowany model dawał lepsze wyniki niż pomiary ważności kosztów.

GILL A. **Optimisation of the technical object maintenance system taking account of risk analysis results**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 420–431, http://dx.doi.org/10.17531/ein.2017.3.13.

The article presents the author's original method of optimisation of the technical object maintenance system taking account of risk analysis results. An original form of the objective function was formulated, in which a risk measure model based on two criteria was used. RBM methods were discussed and technical object maintenance methods/strategies were reviewed, with their most important characteristics pointed out. The process of making maintenance-related decisions is armed with procedures based on a risk valuation pattern. The author's original risk valuation pattern was presented and special cases resulting from the use of such patterns were discussed. Dynamic programming was used to solve the problem of optimisation. The author's original mathematical model of the method of optimisation was developed and presented, and its four-stage calculation algorithm was presented in detail. Based on the collected statistical data on damage, hazard analysis and risk assessment procedures were carried out. Using computer implementation of the optimisation model, an experiment in planning the maintenance of the technical objects examined was carried out and the results of the optimisation experiment were presented.

GILL A. **Optymalizacja systemu obsługi obiektów technicznych z uwzględnieniem wyników analizy ryzyka**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 420–431, http://dx.doi.org/10.17531/ein.2017.3.13.

W artykule przedstawiono autorską metodę optymalizacji systemu obsługi obiektów technicznych z uwzględnieniem wyników analizy ryzyka. Sformułowano oryginalną postać funkcji celu, w której użyto modelu miary ryzyka opartego na dwóch kryteriach. Omówiono metody RBM oraz dokonano przeglądu metod/strategii obsługiwania obiektów technicznych wskazując ich najważniejsze cechy. Podejmowanie decyzji obsługowych uzbrojone jest w procedury oparte na schemacie wartościowania ryzyka. Przedstawiono autorski schemat wartościowania ryzyka i omówiono szczególne przypadki wynikające z użycia takich schematów. Do rozwiązania problemu optymalizacji użyto programowania dynamicznego. Opracowano i przedstawiono autorski matematyczny model metody optymalizacji oraz szczegółowo zaprezentowano jego czteroetapowy algorytm obliczeniowy. Na podstawie zebranych danych statystycznych dotyczących uszkodzeń, przeprowadzono procedury w zakresie analizy zagrożeń i oceny ich ryzyka. Wykorzystując implementację komputerową modelu optymalizacyjnego przeprowadzono eksperyment w zakresie planowania obsług rozpatrywanych obiektów technicznych oraz przedstawiono wyniki eksperymentu optymalizacyjnego.

SUN L, GU X-H, SONG P, DI Y. **A generalized equivalent temperature model in a time-varying environment**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 432–440, http://dx.doi.org/10.17531/ein.2017.3.14.

Accelerated degradation test at high temperature level is a common method to accelerate the degradation of products by elevating temperature, and the obtained degradation data are then used to obtain the estimate of the performance at normal temperature after extrapolating the degradation through accelerating model. However, the normal temperature is ever-changing rather than constant. Therefore, a generalized equivalent temperature model based on power law degradation path is proposed to establish a connection between accelerated degradation data and degradation data at normal temperature. The model takes the equal degradation measure as a principle and the conclusion is demonstrated that the increments of the degradation under the same magnitude, same time and different orders of temperature stresses are same. The result shows that the empirical equivalent temperature model is a special case of the proposed model. The accuracy of the proposed model is finally demonstrated by a case study of nitrile rubber O-rings.

SUN L, GU X-H, SONG P, DI Y. **Uogólniony model temperatury równoważnej w zmiennym w czasie otoczeniu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 432–440, http://dx.doi.org/10.17531/ein.2017.3.14.

Przyspieszone badania degradacji (badania starzeniowe) prowadzone w warunkach wysokiej temperatury stanowią powszechnie stosowaną metodę przyspieszania starzenia produktów poprzez podwyższanie temperatury. Otrzymane w takich badaniach dane degradacyjne wykorzystuje się do szacowania wydajności produktu w temperaturze normalnej na zasadzie ekstrapolacji. Głównym ograniczeniem tej metody jest fakt, że normalna temperatura nie jest stała lecz zmienia się w czasie. Dlatego też, aby skorelować dane z przyspieszonej degradacji z danymi dotyczącymi starzenia w normalnej temperaturze, zaproponowaliśmy uogólniony model temperatury równoważnej oparty na krzywej degradacji opisanej prawem potęgowym. W modelu przyjęto zasadę równego stopnia degradacji i wykazano, że przyrosty degradacji przy tej samej wartości i czasie działania naprężeń termicznych różnego rzędu są takie same. Wyniki pokazują, że empiryczny model temperatury równoważnej jest szczególnym przypadkiem proponowanego przez nas modelu. Trafność opisanego w pracy modelu wykazano na podstawie studium przypadku dotyczącego uszczelek nitrylowych, tzw. oringów.

STĘPIEŃ S, SZAJNAR S, JASZTAL M. **Problems of military aircraft crew's safety in condition of enemy counteraction**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 441–446, http://dx.doi.org/10.17531/ein.2017.3.15.

The presented paper consists outline of the probabilistic method of evaluation of military aircraft crew's safety, which took into consideration enemy counteraction. The specific attention was focused on estimation of durability of ejection seat, which is a means of pilot's emergency escape from aircraft. The basis of the presented model is probability of pilot's danger to life for single sortie caused by enemy. Formulated differentiation equation characterises process of increment of successful sortie number. The equation after transformation into partial differential equation served for establishing of successful sortie distribution function and subsequently for calculation of crew safety indicators.

STĘPIEŃ S, SZAJNAR S, JASZTAL M. **Problemy bezpieczeństwa załogi wojskowego statku powietrznego w warunkach przeciwdziałania przeciwnika**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 441–446, http://dx.doi.org/10.17531/ein.2017.3.15.

W artykule przedstawiono zarys probabilistycznej metody oceny bezpieczeństwa załogi wojskowego statku powietrznego uwzględniającej niszczące działanie przeciwnika. Największą uwagę skupiono na szacowaniu trwałości fotela katapultowego jako środka do awaryjnego opuszczania samolotu przez pilota. Podstawą prezentowanego modelu jest prawdopodobieństwo powstania zagrożenia dla życia pilota w pojedynczym locie statku powietrznego spowodowane przeciwdziałaniem przeciwnika. Sformułowano równanie różnicowe charakteryzujące w ujęciu probabilistycznym proces przyrostu liczby udanych lotów bojowych statku powietrznego. Równanie to po przekształceniu w równanie różnicz-

HE L, LI G, XING L, CHEN Y. **An autonomous multi-sensor satellite system based on multi-agent blackboard model**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 447–458, http://dx.doi.org/10.17531/ein.2017.3.16.

Traditional Earth observation satellite cannot work well in terms of emergencies, environmental uncertainties and scientific events discovery. Therefore, it is of significance to study the new generation of autonomous Earth observation satellite. In order to develop an autonomous satellite system with distributed and coordinated functions, this paper proposes an autonomous satellite system based on distributed multi-agent blackboard model. Multiple agents including functions of pre-processing, planning, scheduling and execution are designed. Agents share information and communicate through a blackboard which stores the task sequence, the action sequence and the satellite status. An adaptive rule-based heuristic scheduling algorithm and a forward search planning algorithm are proposed. The simulation experiments and computational results prove that the system can deal with scientific events discovery, satellite faults, cloud obscuration and emergencies without human intervention, which can greatly enhance the efficiency and reliability of Earth observation satellites. The validity of the proposed model and algorithm is proved.

GUTTEN M, KORENCIAK D, KUCERA M, SEBOK M, OPIELAK M, ZUKOWSKI P, KOLTUNOWICZ T. **Maintenance diagnostics of transformers considering the influence of short-circuit currents during operation**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 459–466, http://dx.doi.org/10.17531/ein.2017.3.17.

Article presents theoretical, simulation and experimental analyses of possible effect of short-circuit forces on the transformer windings. The first part of the article is focused to the establishment and activity radial and axial forces during short circuit. It shows dimensions, direction and of course caused mechanical stress. Equation shows basic dependencies of these mechanical forces created in the transformer windings. The last part of the article is focused on the simulation method which shows the mechanical stress caused by the short-circuit currents on transformer. The paper presents experimental methods of diagnostics for analysis of the short circuit on transformer windings.

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, http://dx.doi.org/10.17531/ein.2017.3.18.

Software reliability modeling has undergone a continuous evolution over the past three decades to adapt to various and ever-changing software testing environments. In existing models, immediate fault removal and fault independency are two basic and commonly used assumptions. Recently, models combining fault detection process (FDP) and fault correction process (FCP) were proposed to alleviate the immediate fault removal assumption. In this paper, we extend such a methodology by proposing a modeling framework for the FDP and FCP incorporating fault dependency. Faults are classified as leading faults and dependent faults and the FCPs for both types of faults are explicitly modeled. Several paired models considering different assumptions for debugging lags are proposed for the combined FDP and FCP. The applicability of the proposed models are illustrated using real testing data. In addition, the optimal software release policy under this framework is studied.

SKRICKIJ V, BOGDEVIČIUS M, ZYGIENĖ R. **Evaluation of the spur gear condition using extended frequency range**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 476–484, http://dx.doi.org/10.17531/ein.2017.3.19.

The paper focuses on working out an algorithm for spur gear condition monitoring, based on the results of numerical simulation. The nonlinear mathematical model has been used for investigation of the dynamic parameters of the cylindrical spur gear with defective teeth. Backlash between gear teeth, backlash in bearings, time-varying mesh stiffness, and variations of the centre distance have been evaluated in the model. Diagnostic parameters suitable for determining the condition of the gears under investigation have been established. Frequency intervals mostly affected by changes in diagnostic parameters under damage have been found. An algorithm for diagnostics based on mathematical modelling, vibro-acoustic, and acoustic emission methods, and wavelet transform has been worked out.

kowe cząstkowe, posłużyło do wyznaczenia funkcji rozkładu udanych lotów bojowych, a następnie wskaźników bezpieczeństwa załogi.

HE L, LI G, XING L, CHEN Y. **Autonomiczny wieloczujnikowy system satelitarny oparty na wieloagentowym modelu tablicowym**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 447–458, http://dx.doi.org/10.17531/ein.2017.3.16.

Tradycyjne satelity obserwacji Ziemi nie nadają się do pracy w sytuacjach kryzysowych, warunkach niepewności środowiskowej oraz w okolicznościach związanych z odkryciami naukowymi. Dlatego też istotne znaczenie ma badanie nowej generacji autonomicznych satelitów obserwacji Ziemi. W celu opracowania autonomicznego systemu satelitarnego o rozproszonych i skoordynowanych funkcjach, w niniejszej pracy zastosowano rozproszony wieloagentowy model tablicowy. Zaprojektowano agenty, w tym funkcje wstępnego przetwarzania, planowania, harmonogramowania i wykonania. Agenty te wymieniają między sobą informacje i komunikują się za pośrednictwem tablicy (ang. blackboard), na której przechowywane są informacje dotyczące sekwencji zadań i działań oraz stanu satelity. Zaproponowano adaptacyjny, regułowy, heurystyczny algorytm harmonogramowania oraz algorytm planowania metodą wyszukiwania w przód. Przeprowadzone eksperymenty symulacyjne oraz wyniki obliczeń dowodzą, że omawiany system sprawdza się w przypadkach odkryć naukowych, awarii satelitarnych, zachmurzenia oraz w sytuacjach kryzysowych nie wymagając interwencji człowieka, co może znacznie zwiększać wydajność i niezawodność satelitów obserwacji Ziemi. W pracy wykazano trafność proponowanego modelu i algorytmów.

GUTTEN M, KORENCIAK D, KUCERA M, SEBOK M, OPIELAK M, ZUKOWSKI P, KOLTUNOWICZ T. **Eksploatacyjne diagnostyki transformatorów uwzględniające prądy zwarcia podczas pracy**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 459–466, http://dx.doi.org/10.17531/ein.2017.3.17.

Artykuł przedstawia teoretyczną, symulacyjną i doświadczalną analizę ewentualnych wpływów prądów zwarciowych na uzwojenia transformatorów. Pierwsza część artykułu koncentruje się na powstaniu i działaniu promieniowych i osiowych sił podczas zwarcia. Przedstawia rozmiary, kierunek i oczywiście powstałe naprężenia mechaniczne. Równanie pokazuje podstawowe zależności sił mechanicznych powstałych w uzwojeniach transformatora. Ostatnia część pracy dotyczy metody symulacji, która przedstawia naprężenia mechaniczne spowodowane przez prądy zwarciowe na transformatorze. Artykuł przedstawia eksperymentalne metody diagnostyki do analizy wpływu zwarcia na uzwojenia transformatora.

PENG R, ZHAI Q. **Modelowanie procesów wykrywania i korekcji błędów oprogramowania z założeniem wzajemnej zależności błędów**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 467–475, http://dx.doi.org/10.17531/ein.2017.3.18.

Modelowanie niezawodności oprogramowania w ciągu ostatnich trzech dekad ulegało ciągłej ewolucji, pozwalającej dostosować je do różnych, stale zmieniających się środowisk testowych. W przypadku istniejących modeli, dwoma podstawowymi i powszechnie stosowanymi założeniami jest natychmiastowe usunięcie błędu oraz brak zależności między błędami. Ostatnio, badacze zaproponowali modele, które łagodzą pierwsze z tych założeń, łącząc proces wykrywania błędów (FDP) z procesem ich korekcji (FCP). W niniejszym artykule, rozszerzono tę metodologię, proponując paradygmat modelowania dla zintegrowanych procesów FDP i FCP uwzględniający zależności między błędami. W paradygmacie tym, błędy klasyfikuje się jako błędy nadrzędne i błędy zależne, a procesy FCP dla obu typów błędów są modelowane oddzielnie. Zaproponowano kilka połączonych w pary modeli rozważających różne założenia dotyczące opóźnień debugowania w procesach łączących detekcję i korekcję błędów. Możliwość zastosowania proponowanych modeli przedstawiono na przykładzie rzeczywistych danych testowych. Dodatkowo badano optymalną politykę aktualizacji oprogramowania, jaką można prowadzić w ramach proponowanego paradygmatu.

SKRICKIJ V, BOGDEVIČIUS M, ZYGIENĖ R. **Ocena stanu przekładni zębatej z wykorzystaniem rozszerzonego zakresu częstotliwości**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 476–484, http://dx.doi.org/10.17531/ein.2017.3.19.

Celem artykułu było opracowanie algorytmu monitorowania stanu przekładni zębatej w oparciu o wyniki symulacji numerycznej. Przedstawiono nieliniowy model matematyczny, który wykorzystano do badania parametrów dynamicznych przekładni zębatej walcowej z uszkodzonymi zębami. Za pomocą przedstawionego modelu oceniano luz pomiędzy zębami przekładni, luz w łożyskach, zmienną w czasie sztywność zazębienia oraz zmiany odległości osi. Ustalono parametry diagnostyczne odpowiednie dla określenia stanu technicznego badanych przekładni. Znaleziono przedziały częstotliwości odpowiadające zmianom parametrów diagnostycznych wynikającymi z uszkodzenia. Opracowano algorytm diagnostyczny oparty na modelowaniu matematycznym, metodach emisji wibroakustycznej i emisji akustycznej oraz transformacie falkowej.

ZHANG S, SUN S, SI S, WANG P. **A decision diagram based reliability evaluation method for multiple phased-mission systems**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 485–492, http://dx.doi.org/10.17531/ein.2017.3.20.

The multiple phased-mission system (MPMS) exists widely in practical engineering, such as aviation, spaceflight and navigation fields. Its distinct characteristic is that the system usually performs multiple missions and each mission consists of different phases. In this paper, we mainly focus on the reliability analysis for MPMS when the components have to accomplish different missions successively. A new modeling method is proposed for MPMS analysis based on the binary decision diagram (BDD) and multi-state multi-valued decision diagram (MMDD). Through this method, different phases of missions are combined with in the whole system by certain merging rules according to the operating time of a common component. Then, the system reliability can be calculated by the common calculation methods of decision diagrams by generating the through. Finally, two case studies are implemented to demonstrate the generation of BDD/MMDD models and the evaluation of system reliability. The experiment results verified the efficiency and accuracy of the proposed modeling methods.

ZHANG S, SUN S, SI S, WANG P. **Metoda oceny niezawodności systemów wielofazowych w oparciu o diagramy decyzyjne**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 485–492, http://dx.doi.org/10.17531/ein.2017.3.20.

Systemy wielofazowe (Multiple Phased-Mission Systems, MPMS), t.j. systemy o wielu zadaniach okresowych są powszechnie stosowane w praktyce inżynieryjnej, np. w lotnictwie, lotach kosmicznych czy nawigacji. Cechą wyróżniającą tego typu systemy jest to, że zazwyczaj wykonują one wiele zadań, z których każde składa się z różnych faz. Głównym tematem poniższej pracy jest analiza niezawodności MPMS dla przypadków, kiedy elementy składowe muszą wykonywać różne misje jedna po drugiej. W artykule zaproponowano nową metodę modelowania dla celów analizy MPMS opartą na koncepcji binarnego diagramu decyzyjnego (binary decision diagram, BDD) oraz wielostanowego wielowartościowego diagramu decyzyjnego (multi-state multi-valued decision diagram, MMDD). Metoda ta polega na łączeniu różnych faz misji w obrębie systemu za pomocą pewnych reguł łączenia wedle czasu pracy wspólnego elementu składowego. Pozwala to na obliczanie niezawodności systemu za pomocą powszechnie stosowanych metod diagramów decyzyjnych poprzez generowanie drzew błędów. W pracy zaprezentowano dwa studia przypadku, które pokazują, w jaki sposób generuje się modele BDD/MMDD oraz ocenia niezawodność systemu. Wyniki eksperymentów wykazały wydajność oraz trafność proponowanych metod modelowania.

Article citation info:
PIETRUCHA-URBANIK K, STUDZIŃSKI A. Case study of failure simulation of pipelines conducted in chosen water supply system. Eksploatacja i Niezawodnosc – Maintenan ce and Reliability 2017; 19 (3): 317–323, http://dx.doi.org/10.17531/ein.2017.3.1.

Katarzyna PIETRUCHA-URBANIK
Andrzej STUDZIŃSKI

CASE STUDY OF FAILURE SIMULATION OF PIPELINES CONDUCTED IN CHOSEN WATER SUPPLY SYSTEM

SYMULACJE AWARII RUROCIĄGÓW PRZEPROWADZONE NA WYBRANEJ SIECI WODOCIĄGOWEJ – ANALIZA PRZYPADKU

The main goal of this work is to simulate the failure of water pipe network, using the hydraulic model of the network created through Epanet 2 program. The model includes the cooperation of the second stage pumping station with the expansion tanks located in the network. Based on these parameters, the simulation operation of water supply network was performed, as well as failure simulation on the basis of closing some sections of water pipe network. Failure analysis allowed to perform characteristics of the water supply system including pressure changes that occur in the network during failure simulation.

***Keywords**: water supply, failure, hydraulic model, risk, reliability.*

Głównym celem pracy jest przeprowadzenie symulacji awarii przewodów przewodów wodociągowych, za pomocą stworzonego modelu hydraulicznego sieci z wykorzystaniem programu Epanet 2. W modelu zawarta jest współpraca pompowni drugiego stopnia ze zbiornikami wyrównawczymi znajdującymi się na sieci. Bazując na tych parametrach przeprowadzono symulację pracy sieci wodociągowej, a dzięki wyłączeniem odcinków przewodów wodociągowych symulację awarii. Analizowanie awarii pozwoliło na scharakteryzowanie dostawy wody do odbiorców w tym zmian, jakie występują na sieci podczas symulowania awarii.

***Słowa kluczowe**: wodociąg, awaryjność, model hydrauliczny, ryzyko, niezawodność.*

1. Introduction

In the Safe Drinking Water Act (SDWA) passed in 1974 by Congress of the United States the strict regulation for water supply was presented. Also water recipients can get from the Environmental Protection Agency a Consumer Confidence Report which precisely describes the water network functioning, along with potential risks in drinking water network. In spite of incorporating many reports, regulations and standards that public drinking water systems must follow, still many failures occur causing serious effects for water consumers. The main problem that may arise in the operation of water supply systems are breaks in water supply and change in water quality in case of failure when the standard requirements are not met [4, 9]. Risk that users of the water supply system take in such case is the lack of water supply or receiving water with insufficient pressure and inadequate quality [25]. The risk that such undesirable events occur is associated with consumer dissatisfaction [11]. The risk concerns not only the water consumers but also the water producers. From the producers point of view it involves disturbances or interruptions in water production or distribution, which, in turn, causes the financial losses caused by the unsold water, penalties and compensations paid to water recipients [24]. Such situations are caused by the water supply system failures, especially by failures in the main network which have larger consequences and cover larger area, while failures of distributional pipes are important to customers directly supplied from given pipes and the nearest area. To prevent such problems and improve the quality of services provided by water companies, risk analysis should be conducted by hydraulic simulating failures in water pipes [1, 2] and through implementation of intelligent monitoring of water supply system [26, 27]. Currently the requirements relating to water supply increase, therefore the water suppliers should reduce the risk associated with lack of water supply or supply poor quality water [14]. Therefore risk analysis and assessment in municipal systems is commonly used [3, 7, 16, 23]. Determination of the risk associated with failure rate of water supply network consists of several phases, such as the determination of water supply system type, the designation of the failure rate limit, the determination of the difficulty and type of repairs, determining the protection barriers and the risk levels [17].

Risk is generally defined as the expected value of the losses and can be presented by the formula (1) [18]:

$$R = P \cdot C \tag{1}$$

where P is the probability of failure occurrence and C are consequences associated with the probability P.

Failure causes of the municipal infrastructure involve such errors as design errors, both in the hydraulic calculations, errors of exploitation character, e.g. no security resistance, lack of insulation, incorrect compensation, incorrect execution of connections, improper laying of pipes and operating errors such as inadequate strategy of repairs, tech-

nical wear, exploiters errors and aggressive or sulfate media influence [3, 12, 13, 19].

Therefore for modelling water mains are used different kinds of programs that are very helpful in solving problems related to the design and operation of water distribution systems [6, 21].

The basic role played by computer models is the hydraulic analysis of water distribution system operation. This analysis may include not only changes in water pressure in the network but also water quality changes in the network via the information about the spread of disinfectant, the cost of the electricity used to pump water [22]. To facilitate the analysis of consequences of failure occurrence and thus the scope and duration or size of restrictions in water supply the hydraulic model mapping the network operation was developed in the program Epanet 2. This model shows the current state of the water demand, together with the current water distribution in the analysed system supplying water for about 180 thousand of residents. In this model also the cooperation of the second degree pumping station with the expansion tanks was presented.

The analysis referred to comparison of the network working without failure for 24 hours with the operation of the network when the failure occurs. The proposed model allowed to determine to perform analysis of operation of the water supply network. The scope of work includes compiling information about the water supply system and making the hydraulic model of water supply network, performing hydraulic simulation of pipeline failure, presentation and summarizing the results of the simulation.

2. Case study of failures in water network based on operational data

2.1. Preliminary assessment of pipeline technical state

The analysed water network is a ring system formed by four mains with total length of 49.8 km. The total length of operated water supply system amounted to 902.8 km in 2016. Figures 1 and 2 show the age and material structure of the considered water supply network.

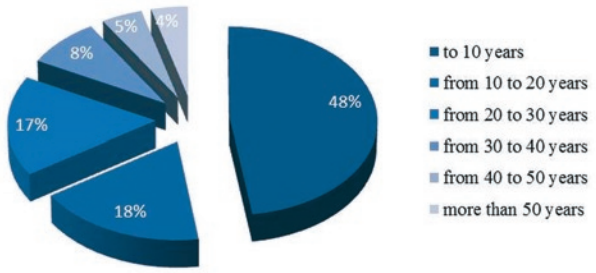


Fig. 1. The age structure of the water network - state for 2016, in %

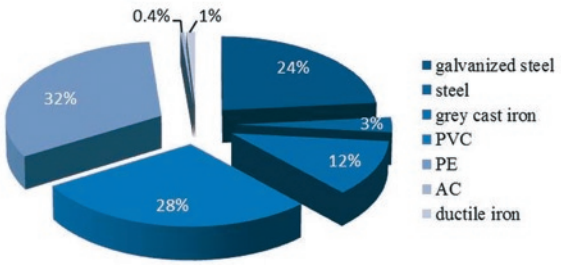


Fig. 2. The material structure of the water network - state for 2016, in %

Data on failures connected with failure time removal in the water pipes in the years 2000-2016 are presented in Fig. 3.

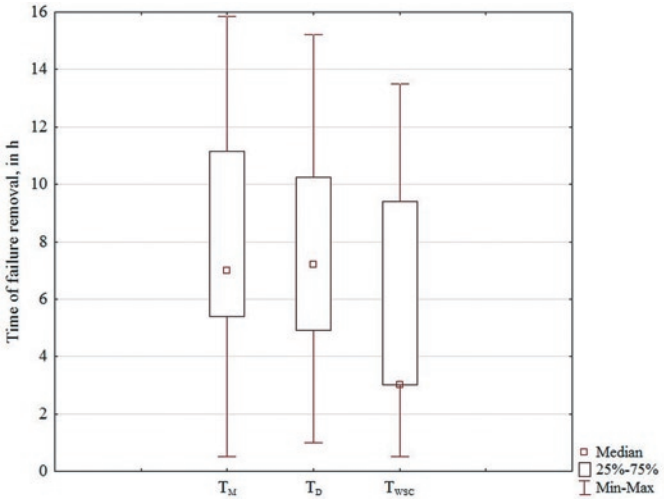


Fig. 3. Time of failure removal in the water supply pipelines in the years 2000-2016 for mains - T_M, for distributional pipes - T_D, and for water supply connections - T_{wsc}

In the Figure 4 the number of failures expressed in percentage depending on the pipe material was presented.

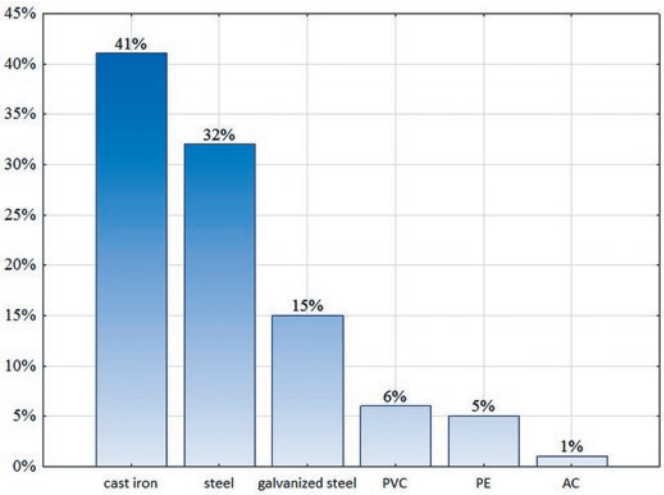


Fig. 4. Failures of water network in the years 2000-2016 with division into failure occurrence and pipe material, in %

2.2. Water network failure

The unit failure rate λ_i, with division into water pipe type, material and type of failure was calculated according to the formula (2) [8]:

$$\lambda_i = k_i / (l_i \cdot \Delta t)$$

where k_i is the total number of failures in one year in a given type of network, l_i is the length of a given type of network in km, Δt is the considered period, one year, and i is the type of water network or type of failure.

2.3. Discussion of results

The average values of the failure rate for different types of network are:

- for mains $\lambda_M = 1.03\ a^{-1} \cdot km^{-1}$,
- for distributional pipes $\lambda_D = 0.34\ a^{-1} \cdot km^{-1}$,

- for water supply connections $\lambda_{wsc} = 0.32\ a^{-1}\cdot km^{-1}$,
- for the whole water network $\lambda_t = 0.56\ a^{-1}\cdot km^{-1}$.

The failure rate for each type of water supply system is shown in the Figure 5.

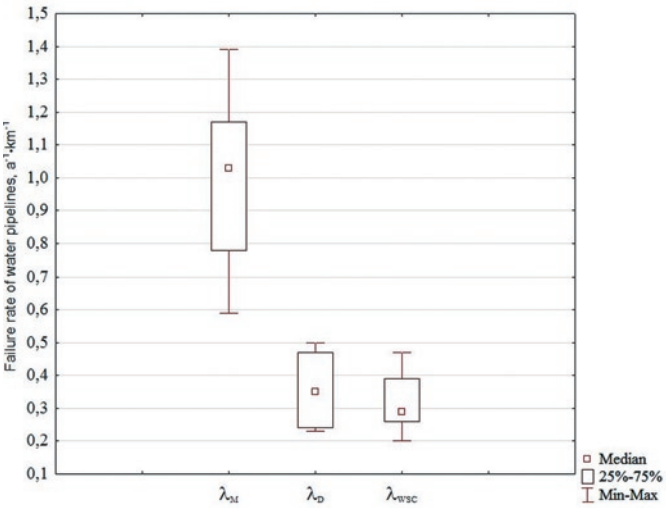


Fig. 5. The failure rate of the mains - λ_M, distributional pipes - λ_D, and water supply connections - λ_{wsc}

Analysing the failure rate it can be stated that in the years 2000-2016 the number of failures and thus the intensity of failures steadily decreases, which in turn can be caused by modernization of the network in recent years and the implementation of new pipes made of more durable materials that are less prone to failures. It can be seen that the highest failure rate is for the mains, it is caused by the pipe age, also by the fact that these pipes are made of iron or steel, which are more prone to failure than PE pipes, for which most old pipes are replaced. The failure rate for the mains should not exceed $0.3\ a^{-1}\cdot km^{-1}$ according to risk criteria given in [15], so it can be concluded that the mains require renovation or replacement. The distributional network and water supply connections do not exceed the limit values for this indicator, which is for the distributional network $0.5\ a^{-1}\cdot km^{-1}$ and for water supply connections $1.0\ a^{-1}\cdot km^{-1}$. Taking into account the material structure of network pipes, the highest failure rate have water pipes made of cast iron ($\lambda_{cast\ iron} = 1.5\ a^{-1}\cdot km^{-1}$) and steel ($\lambda_{steel} = 1.53\ a^{-1}\cdot km^{-1}$). Currently used plastic materials like PVC and PE indicate lower failure rate, respectively, $\lambda_{PVC} = 0.03\ a^{-1}\cdot km^{-1}$ and $\lambda_{PE} = 0.06\ a^{-1}\cdot km^{-1}$. Increased failure rate for specific materials determines the cause of the failure, the highest rates were recorded for the failures caused by corrosion and cracks, the median of mentioned causes amounted to, respectively, $\lambda_{corrosion} = 0.36\ a^{-1}\cdot km^{-1}$ and $\lambda_{cracks} = 0.26\ a^{-1}\cdot km^{-1}$.

3. Research methodology

3.1. Characteristics of software used for the simulation of the water supply system

One of the recommended program used for water supply system modelling is program called Epanet 2.0, created by Lewis A. R. for the U.S. Environmental Protection Agency. It is used to design and control the operation of water supply systems. In this program, as in a real network, the following elements can be found: intake, pumps, nodes, different kinds of valves and tanks. The program also allows to track the flow in each pipe or the amount of pressure in the individual node [20].

Work with the program Epanet 2.0 begins by drawing the water supply network model, then the characteristics of elements used to create the tested model of examined water network should be made. They are, among others:

- reservoir (water source for the modelled network). Input data include: determining the location of a reservoir in the diagram, the elevation of the water table above sea level, the quality of water flowing into the network,
- joint (network node which starts and ends pipe). Input data: the location of the joint in the scheme, the elevation of the joint above sea level, the time distribution of water demand in the given node, the number of different categories of water consumption defined for the node, coefficient reducing the flow associated with the applied flange assigned to the node, quality of water entering the network at this point,
- pipe (pipe connecting two nodes). Input data: length, diameter, absolute roughness of the pipe, the loss factor,
- tank. Input data: determining the position of the tank in the scheme, the position of the tank bottom above sea level, the initial water level, the minimum and maximum determine the level of water in the tank, diameter, minimum volume, volume curve, mix model, initial quality and source,
- pump. Input data: coordinates, identifier of a node in which the suction line begins and ends, the characteristics of the pump flow, power, velocity, technical data of the pump, the characteristics of efficiency, energy price - the price for energy calculated implicitly for kW/h, the cost of energy for one day.

In this study only the results of modelling the tested water supply system based on the hydraulic and time characteristics are presented.

3.2. Hydraulics of water supply network

To help minimize the consequences of the failure of the water supply pipeline and hence range, duration and size of interruptions in water supply the hydraulic model was developed in the Epanet 2, mapping the network operation. It allows to simulate failure of individual network sections defining the scope of the impact of the section exclusion on the network operation.

It should be noted that the water supply system constitutes a set of interrelated elements which work affects other elements. The parameters which describe the operation of these elements are, among others, the flow rate, pressure, flow resistance. The consequences of such structure is the need for modelling the entire water supply system [5, 9, 10].

Before performing the simulation the water demand was updated for each node of the model. Pipes of the model are assigned to individual streets along which they are laid, in accordance with the updated map obtained from the water company. In the model, a number of water meters in the street and the number of inhabitants supplied from the given pipe, were determined. The value of the individual water demand was established after taking into account the work of tanks, reservoirs and therefore cooperation with the pumping station, readings of water meters and water level fluctuations in the expansion tanks. It allowed to determinate the daily and hourly water demand.

In case of high buildings which are supplied by hydrophores the data were obtained from the water company. High impact on the water demand have the industrial plants, which were also included in the analysis after receiving the detailed data from the plants.

The presented simulation covers one day of largest consumption in the water network, 24 hours beginning from 6 a.m. With the model it is possible to perform the analysis of the consequences of network pipe failure by comparing the pressure in the nodes of the network during normal network operation and during failure occurrence.

Before performing calculations on the network model the tare was conducted which involved determining the characteristics of the sec-

ond stage pumping station to illustrate real city's water supply and to compare the determined pressure in the network to the real conditions, taking into account the information obtained from network monitoring. To calculate the hydraulic losses the Darcy-Weisbach formula was used.

3.3. Determination of the risk associated with the failure rate of the tested water network using two-parameter method

Using the calculated failure rate the risk of failure in water supply network can be calculated. For calculating risk the two-parameter method was proposed according to the equation 1.

The probability of failure P can be determined by multiplying the failure rate, the pipe length and pipe renewal time (Eq. 3):

$$P = \lambda \cdot l \cdot t \tag{3}$$

where λ is the failure rate for each pipe, ($a^{-1} \cdot km^{-1}$), L is the length of the examined segment of pipe (km) and t is renewal time of examined segment of pipe, (a^{-1}).

The consequence of failure is the number of inhabitants without the access to water, so the risk was determined according to Eq. 1 by multiplying the consequence and failure probability.

The limit values of risk have been adopted on the basis of the analysis of the operational data of the analysed water supply system. The limit value of tolerable risk was determined by multiplying the probability of failure falling on one segment of the distribution network and the number of residents cut off from water supply per one section of the network, respectively R = 0.085. The limit value for controlled risk has been adopted for the main network and amounted to R = 0.62.

4. Results of the model analysis of water supply network in the Epanet program

The first step was to analyse all the studied sections and turn off each of them. Examples of the results of the pressure head difference in the nodes were shown in Table 1. Due to the large amount of results only part of them was presented and selected hours with the highest demand for water in which there was the highest pressure drop.

In most cases, the pressure head difference in the examined nodes before and after the failure is insignificant, which indicates that the failure will affect only the residents directly supplied from the pipeline, it is caused by the fact that the water supply network is constructed as a ring . Because the analysis contain large amount of results, for nearly 300 pipes, only some failures were selected to present in order to indicate their consequences, as shown in Table 2.

In a detailed manner the simulation results for two pipes located in various parts of the city whose exclusion causes a significant decrease in pressure at the site of the failure and its surroundings, were presented. Such pipes were chosen whose exclusion caused head drop of more than 1.0 mH_2O.

Results during operation with and without failure are presented on the Figures 6-9.

In the presented sections of pipes the risk value exceeded the tolerable level. In the section no. 1571 with a length of 104.4 m and a diameter of 300 mm, the failure causes that 5940 inhabitants are without water supply for R = 0.4885.

The largest pressure difference in this pipe, before and after failure, occurred in the node 157 at 6.00 and 10.00, and reached 5.01 m, from the value of 34.72 mH_2O to 29.71 mH_2O at 6.00. On the other hand, in the section 98 the biggest pressure difference before and after failure occurred in the node 102 at 6.00 and amounted to 7.90

Table 1. Part of calculations obtained by modelling each node during operation without failure and during failure in selected hours of highest water demand generated by using Epanet

State before failure		State after failure		State before failure		State after failure	
Time [h]	Pressure head [m]	Time [h]	Pressure head [m]	Time [h]	Pressure head [m]	Time [h]	Pressure head [m]
Table of time series – Nod 503				Tabela serii czasu – Węzeł 750			
00:00	32.30	00:00	28.16	00:00	29.73	00:00	26.05
00:01	32.27	00:01	28.38	00:01	29.74	00:01	26.28
00:02	32.07	00:02	27.99	00:02	29.51	00:02	25.88
00:03	32.03	00:03	28.14	00:03	29.49	00:03	26.04
00:04	31.83	00:04	27.75	00:04	29.26	00:04	25.65
00:05	31.80	00:05	27.98	00:05	29.27	00:05	25.88
00:06	31.76	00:06	28.13	00:06	29.26	00:06	26.03
00:07	31.86	00:07	28.69	00:07	29.42	00:07	26.61
⋮							
Table of time series – Nod 1229				Table of time series – Nod 750			
00:00	33.05	00:00	29.76	00:00	22.58	00:00	15.00
00:01	32.96	00:01	29.54	00:01	22.60	00:01	15.48
00:02	32.75	00:02	29.25	00:02	22.36	00:02	14.90
00:03	32.85	00:03	29.12	00:03	22.36	00:03	15.24
00:04	32.68	00:04	28.88	00:04	22.12	00:04	14.66
00:05	32.69	00:05	28.74	00:05	22.14	00:05	15.13
00:06	32.64	00:06	28.60	00:06	22.13	00:06	15.46
00:07	32.73	00:07	28.55	00:07	22.31	00:07	16.49

Table 2. Part of the analysis of the sections of distribution network included in the hydraulic model along with the risk calculated by means of the two-parameter method

L.p.	No of section in Epanet	Street name	Diameter [mm]	Lenght [km]	LM [-]	Risk value
	Remarks about pressure in the nodes involved in the section with failure					
24	1402	Baczyńskiego	150	0.4037	67	0.0043
	The highest increase in pressure occurred in the node 313 at 12.00 a.m. by 0.46 m and the highest drop in pressure at node 1308 at 12.00 a.m. by 1.81 m.					
69	1682	Kościuszki	300	0.0802	89	0.0073
	In the node 24 changes in pressure of several cm. Node 1586 pressure drop at 10.00 a.m. of 1.83 m. Node 1587 pressure drop at 10.00 a.m. of about 1.71 m.					
105	1577	Żółkiewskiego	300	0.6299	100	0.0082
	Node 1487 pressure drop at 1.00 p.m. of 1.66 m. Node 164 increase in pressure at 1.00 p.m. of about 1.43 m.					
124	709	Hetmańska	300	0.116	80	0.0066
	Node 650 increase in pressure at 9.00 a.m. of 1.16 m. Node 391 pressure drop at 2.00 p.m. of about 1.86 m.					
221	98	Łukasiewicza	200	0.214	386	0.264
	Node 102 pressure drop at 6.00 a.m. of about 7.90 m. Node 7 increase in pressure at 6.00 a.m. of 0.32 m.					
239	21	Kochanowskiego	300	0.2056	302	0.0234
	Node 350 increase in pressure at 6.00 a.m. of 0.42 m. Node 1485 pressure drop at 6.00 a.m. of about 5.09 m.					

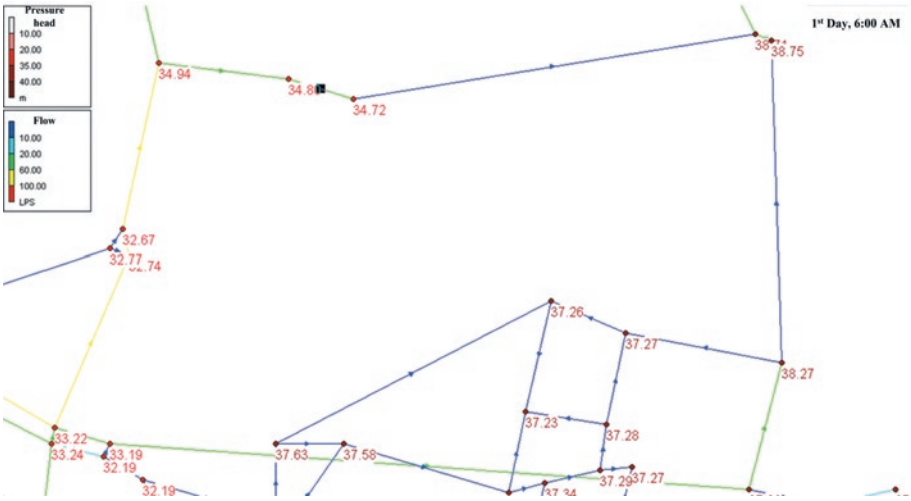


Fig. 6. Simulation of network state before failure - pipe section no 1571

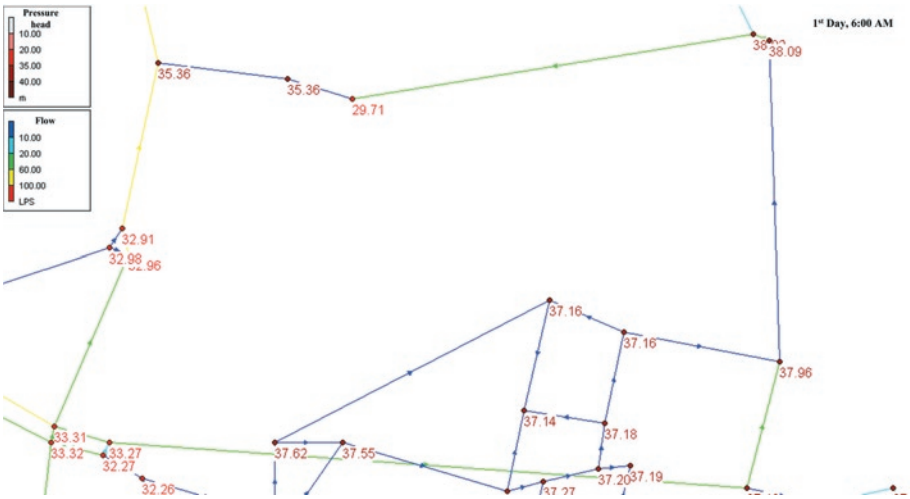


Fig. 7. Simulation of network state after failure - pipe section no 1571

mH_2O from the value of 26.66 to 18.76 mH_2O. Pipe no. 98 supplies water to approximately 386 inhabitants and is characterized by the risk equal to 0.1123. For 23% of examined pipes the risk value exceeds the tolerable risk.

The risk of the tested water supply system takes a small value because during failure of individual water pipes, water is cut off only to the recipients being supplied from those pipes, however, those values are quite large reaching, in extreme cases, several thousand residents without water supply. After detailed analysis of the hydraulic model, it was found that the pressure in the nearest area does not fall enough to cause problems with water supply. A major impact on this situation has also a ring structure included in the model and oversizing of the water supply network.

5. Conclusions and perspectives

The high degree of independence of the network from the failure of individual pipes is probably due to significant oversizing, as well as a significant decrease in water demand, which is now less than 50% of demand in the early nineties of last century. This concerns for example all Polish urban water supply systems developed at least since the sixties, but not in these water systems, which were built since the eighties of the twentieth century.

The performed research was focused on depiction of the hydraulic network, recipients' failure nuisance and real data operation of the exemplary case study water network, during normal and failure event operation. The proposed methodology can be implemented in functioning assessment of any network, dealing with large number of different variables and ob-

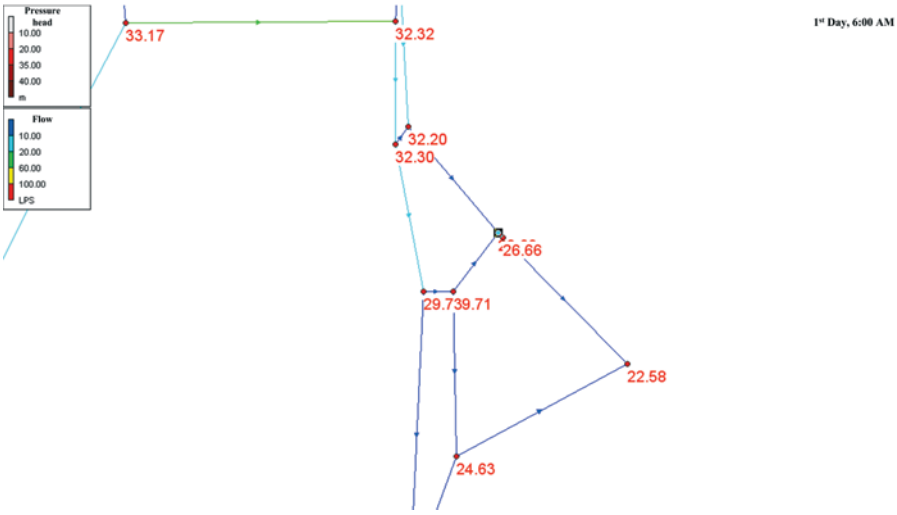


Fig. 8. Simulation of network state before failure - pipe section no 98

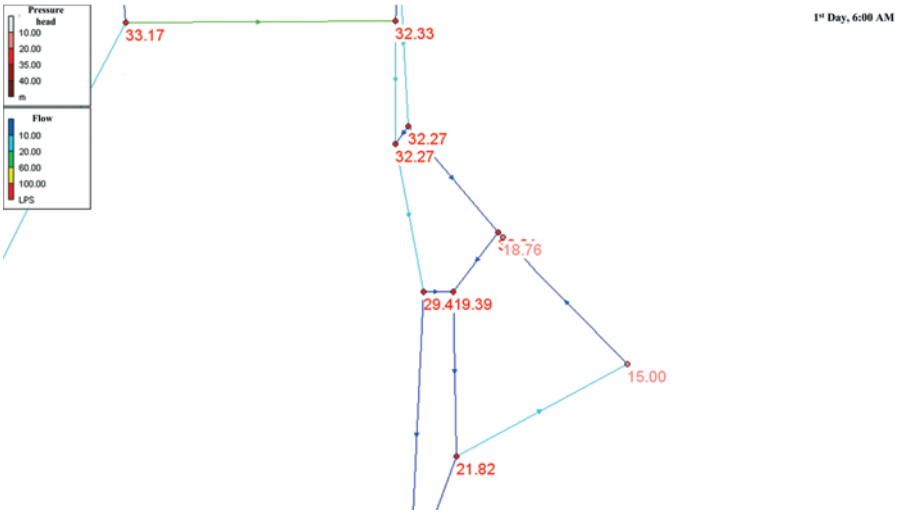


Fig. 9. Simulation of network state after failure - pipe section no 98

tained network information, as to simulate different scenarios and optimize the system. The program used in the analysis does not include cost analysis, but this factors were not used in the consideration of problem with reliable water supplying to recipients. It is worth to underline the possibility of the program to establish the demand fluctuations at each node, as a result of failure occurrence, along with place determination of water supply objects as for example tanks.

A major limitation of the method is to have the revised hydraulic model of water supply network, the construction of which is time-consuming, requires a series of data, which many water supply companies does not possess, as diameter and absolute roughness of pipes built in the first half of the twentieth century, and finally the need to calibrate the model. In practice, only a few waterworks have verified the hydraulic models that can be used in the presented method.

The performed analysis can be an important method to point out the segments of water pipe network which should be modernized in the first place because of their major importance for the recipients, in future research supported by different software through significant performance indicators. Such segments cause the most noticeable for water consumers losses - hence the need for a particular focus on their performance, including qualifying for the reconstruction or renovation. The presented method does not cover all factors influencing the decision-making processes in the activities of the operating water pipes, an outstanding example is the earlier pipe reconstruction cables, despite satisfactory technical state technical, due to the road reconstruction. However this method indicates, however, courses of the water supply system operator in order to achieve the best technical result in the assumed operating conditions of the system.

The presented method and the results are the basis for further studies on risk in quantitative terms, and can be used in the operating practice of water company exploiting the tested water supply system, both with inclusion of internal, external, and environment factors of water supply functioning. Issues outside operation will be covered by different criteria of crucial variables as for example backfilled soil or installation place, through implementation of multi-criteria methodologies and alternative machine learning methods, will constitute the important tool for maintenance support. In future research, the simulation of water network failure will be performed along with cost and economic analysis, based on the past event experience.

References

1. Bene JG, Selek I. Water network operational optimization: Utilizing symmetries in combinatorial problems by dynamic programming. Periodica Polytechnica Civil Engineering 2012; 1: 51-61, https://doi.org/10.3311/pp.ci.2012-1.06.
2. Hallmann C, Suhl L. Optimizing Water Tanks in Water Distribution Systems by Combining Network Reduction, Mathematical Optimization and Hydraulic Simulation. OR Spectrum 2016; 38: 577-595, https://doi.org/10.1007/s00291-015-0403-1.
3. Hotloś H. Ilościowa ocena wpływu wybranych czynników na parametry i koszty eksploatacji sieci wodociągowych. Wrocław: Wroclaw University of Technology Publishing House, 2007.
4. Karamouz M, Moridi A, Nazif S. Urban Water Engineering and Management. New York: Taylor and Francis Group, 2010, https://doi.org/10.1201/b15857.
5. Knapik K. Czasowo-przestrzenna symulacja działania systemu dystrybucji wody. Monografia. Kraków: Wydawnictwo Politechniki Krakowskiej, 1989.
6. Knapik K. Dynamiczne modele w badaniach sieci wodociągowych. Monografia. Kraków: Wydawnictwo Politechniki Krakowskiej, 2000.
7. Kutyłowska M. Modelling of Failure Rate of Water-pipe Networks. Periodica Polytechnica Civil Engineering 2015; 1: 37-43, https://doi.org/10.3311/PPci.7541.

8. Kwietniewski M, Roman M, Kłoss-Trębaczkiewicz H. Niezawodność wodociągów i kanalizacji. Warszawa: Arkady, 1993.
9. Mays LW. Water transmission and distribution. Denver: American Water Works Association, 2010.
10. Mielcarzewicz EW. Obliczanie systemów zaopatrzenia w wodę. Warszawa: Arkady 2000.
11. Nowacka A, Wlodarczyk-Makula M, Tchorzewska-Cieslak B, Rak J. The ability to remove the priority PAHs from water during coagulation process including risk assessment. Desalination and Water Treatment 2016; 3: 1297-1309, https://doi.org/10.1080/19443994.2015.1030108.
12. Ondrejka Harbulakova V, Estokova A, Stevulova N, Luptakova A. Different aggressive media influence related to selected characteristics of concrete composites investigation. International Journal of Energy and Environmental Engineering 2014; 5 (2-3): 1-6, https://doi.org/10.1007/s40095-014-0082-8.
13. Ondrejka-Harbulakova V, Purcz P, Estokova A, Luptakova A, Repka M. Using a Statistical Method for the Concrete Deterioration Assessment in Sulphate Environment. Chemical Engineering Transaction 2015; 43: 2221-2226.
14. Pietrucha-Urbanik K. Failure Analysis and Assessment on the Exemplary Water Supply Network, Engineering Failure Analysis 2015; 57: 137-142, https://doi.org/10.1016/j.engfailanal.2015.07.036.
15. Pietrucha-Urbanik K. Failure Prediction in Water Supply System – Current Issues, Theory and Engineering of Complex Systems and Dependability. In: Advances in Intelligent Systems and Computing, Eds. Zamojski W, Mazurkiewicz J, Sugier J, Walkowiak T, Kacprzyk J. 365, pp. 351-358. Springer International Publishing, Switzerland 2015.
16. Pilch R, Szybka J, Tuszyńska A. Application of factoring and time-space simulation methods for assessment of the reliability of water-pipe networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16 (2): 253-258.
17. Rak J. Metoda planowania remontów sieci wodociągowej na przykładzie miasta Krosna. Czasopismo Inżynierii Lądowej, Środowiska i Architektury - Journal of Civil Engineering, Environment and Architecture. JCEEA 2014; 1: 225-232, https://doi.org/10.7862/rb.2014.15.
18. Rak JR. Some aspects of risk management in waterworks. Ochrona Srodowiska 2007; 29(4): 61-64.
19. Rak J, Kwietniewski M. Niezawodność infrastruktury wodociągowej i kanalizacyjnej w Polsce. Warszawa: Komitet Inżynierii Lądowej i Wodnej PAN, 2010.
20. Rossman LA. Epanet 2 User's Manual, Water Supply and Water Resources Division. Cincinnati: United States Environmental Protection Agency, 2000.
21. Studziński A.: Ryzyko awarii magistrali wodociągowej Szczepańcowa" w Krośnie. Instal 2011; 11: 58-62.
22. Studziński A, Kobylarz J. Jakościowa analiza ryzyka awarii przewodów wodociągowych wodociągu grupowego. Czasopismo Inżynierii Lądowej, Środowiska i Architektury – Journal of Civil Engineering, Environment and Architecture. JCEEA 2014; 1: 311-321, https://doi.org/10.7862/rb.2014.21.
23. Studziński A., Pietrucha-Urbanik K. Preventive maintenance and reliability of water supply system elements. Czasopismo Inżynierii Lądowej, Środowiska i Architektury – Journal of Civil Engineering, Environment and Architecture. JCEEA 2015; 3: 429-436, https://doi.org/10.7862/rb.2015.126.
24. Tchórzewska-Cieslak B, Szpak D. Propozycja metody analizy i oceny bezpieczeństwa dostawy. Ochrona Środowiska 2015; 3: 43-47.
25. Wieczysty A. Niezawodność miejskich systemów zaopatrzenia w wodę. Kraków: Wydawnictwo Politechniki Krakowskiej, 1993.
26. Wyczółkowski R, Matysiak G. The development of an intelligent monitoring system of a local water supply network. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; 42 (2): 71-75.
27. Wyczółkowski R. Intelligent monitoring of local water supply system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2008; 37 (1): 33-36.

Katarzyna PIETRUCHA-URBANIK
Andrzej STUDZIŃSKI
Department of Water Supply and Sewerage Systems
Faculty of Civil, Environmental Engineering and Architecture
Rzeszow University of Technology
al. Powstancow Warszawy, 35-959 Rzeszow, Poland
E-mails: kpiet@prz.edu.pl, astud@prz.edu.pl

Article citation info:
GODZIMIRSKI J, ROŚKOWICZ M, KOMOREK A. Investigation of ballistic resistance of adhesive bonded multi-layer structures. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 324–330, http://dx.doi.org/10.17531/ein.2017.3.2.

Jan GODZIMIRSKI
Marek ROŚKOWICZ
Andrzej KOMOREK

INVESTIGATION OF BALLISTIC RESISTANCE OF ADHESIVE BONDED MULTI-LAYER STRUCTURES

BADANIA ODPORNOŚCI NA PRZEBICIE KLEJONYCH STRUKTUR WIELOWARSTWOWYCH*

The aim of the study was to evaluate the possibility of using multilayer structures for light armour, in which the elements of the ceramic type SiC and Al_2O_3 were bonding by adhesive joins with antiballistic polyethylene and aramid fabrics. Ballistic resistance shells was determined using for this purpose the test stand constructed on the basis of helium gun and high-speed camera. It has been shown that the laminated structure composed of thin metal layers and aramid fabrics as well as polymer composites made of aramid fabric has lower ballistic resistance than loose fabric packs. It was also demonstrated the functionality of the use of the ceramic component as a separate package, which the ceramic plates are adhesive bonded between the two layers of sheet metal. There is also evidence that fabrics composed of thin layers of material poorly connected with each other, should not be adhesively bonded to the ceramic. It proposed the preparation of specimens, which best reconstruct the load of ceramic plates adhesive bonded to fabric, which are made of lightweight bulletproof vests and ballistic panels.

***Keywords**: adhesive joint, multi-layer materials, light-weight ballistic armour.*

Celem badań była ocena możliwości zastosowania opracowanych struktur wielowarstwowych do wytwarzania lekkich pancerzy, mogących znaleźć zastosowanie jako kuloodporne osłony balistyczne śmigłowców bojowych i innego lekkiego sprzętu wojskowego. Istotnym celem badań było również sprawdzenie możliwości łączenia metodą klejenia ceramiki typu SiC i Al_2O_3 z antybalistycznymi tkaninami polietylenowymi i aramidowymi. Odporność na przebicie była badana z wykorzystaniem stanowiska z działem helowym i ultraszybką kamerą. W wyniku przeprowadzonych badań stwierdzono, iż pakiety luźnych tkanin aramidowych charakteryzuje większa odporność na przebicie w porównaniu z kompozytami polimerowymi wytworzonymi z takich tkanin oraz kompozytami warstwowymi złożonymi z cienkich warstw metalowych i tkanin aramidowych. Ponadto zauważono, że warstwy ceramiczne znacznie zwiększają odporność na przebicie osłon balistycznych i ich stosowanie w takich osłonach wydaje się niezbędne oraz, że wklejenie płytek ceramicznych pomiędzy dwie cienkie blachy ze stopu aluminium i nie łączenie ich bezpośrednio z tkaninami aramidowymi zapewnia w przypadku uderzenia pociskiem zniszczenie małej powierzchni warstwy ceramicznej, a więc w małym stopniu zmniejsza właściwości ochronne osłony.

***Słowa kluczowe**: połączenie klejowe, materiał wielowarstwowy, lekka osłona balistyczna.*

1. Introduction

Contemporary light ballistic shields are usually multilayered structures of low density [2, 6, 9, 10]. In the past, the shield used to be a monolithic plate made with high strength steels or titanium alloys. Recently it have been observed a prevailing tendency to use shields which ensure the best possible ballistic protection with the lowest possible mass, at the same time. In order to achieve the above-mentioned effect, manufacturers use multilayer structures, made with various materials. The external part, usually of high strength and of proper hardness (e.g. a layer of ceramics) should be able to withhold or at least deform a penetrator (bullet), which facilitates its "capture" and ensures protection against further perforation of the external layers, most commonly composed of products based on aramid or polythene fibres. Ceramic layers of light ballistic shields are commonly produced with Aluminium oxide (Al_2O_3), Silicon carbide (SiC) and Boron carbide (B_4C) [6, 8]. There is also a large number of investigations over the application of Silicon nitride (Si_3N_4), Titanium boride (TiB_2), Aluminium nitride (AlN), sialons (SiAlON), glasses [4, 6], and ceramic composites reinforced with metal or intermetallic phases [3]. Ceramic layers are most often joined with other layers absorbing the energy, by the adhesive bonding method [1]. In literature the stress is put on the role of adhesive joints in the shaping of protective properties of multilayer armours [5, 7]. For instance the research [7] indicated that for a twin-layer amour (Aluminium oxide/aluminium), there is an optimum thickness of the adhesive layer (0.3 mm), where the effectiveness of the armour is the highest. The most commonly used adhesives are cyanoacrylic ones or adhesives based on epoxide resins.

The aim of this investigation is to assess the possibilities of applying the designed multilayer structures for the manufacture of light armour which may be exploited as bulletproof ballistic shields of combat helicopters and other light military equipment. A crucial part of the investigation is also to check the possibilities of bonding SiC and Al_2O_3 type ceramics with anti-ballistic polythene and aramid fabrics, by means of the adhesive bonding method.

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

2. Methods

2.1. Ballistic resistance of multi-layer structures - object and method of investigation

While taking into account low density of the tested multilayere armours, it was decided that they will be made on the basis of aluminium alloy sheets AW 2024T3 and four aramid fabrics, which differ in their structure and basis weight (g/m^2). Moreover, an investigation are conducted, where the composites were protected by the ceramic layer, prepared on the basis of Silicon carbide SiC.

In order to prepare the specimens, 0.3 mm thick aluminium alloy sheets are used as well as aramid fabrics Microflex, CT 709, T 750 and XPS102. The ceramic layer was composed of plates, whose shape was the right prism, and the base was a regular hexagon (the length of the inscribed circle equalled 20.2 mm and the height was 4.2 mm). An epoxy adhesive Epidian 57 is used, cured by Z1 curing agent to prepare the adhesive bonding structures. The fact that the adhesive is used and not saturant is connected with the fact that the exploited fabrics, in practice, cannot be oversaturated. Moreover, the saturants used for oversaturating the fabrics are characterized with worse adhesion to metals, compared to adhesives.

The surfaces of the aluminium alloy sheets were prepared for bonding by rubbing with an abrasive cloth of granulation 80 and rinsing with extraction gasoline. The attempts of sanding thin aluminium sheets resulted in plastic deformations, which caused that the whole surface could not be bonded. The adhesive layers of the specimens were initially cured in ambient temperature, with the pressure equalling 0.05 MPa for 24 hours, and then for 6 hours at the temperature of 60°C. The specimens of 150x250 mm, prepared in such a way, were later used for cutting specimens with the "water jet" method. The specimens' shape enabled to fix them in a special gripping frame (Fig. 1), placed directly opposite a helium gas gun (Fig. 1). The ballistic resistance was tested by shooting at the specimens with steel bullets of 8 mm in diameter. During the firing at composite specimens, the trajectory of the bullets was registering by means of a camera for fast-changing phenomena (Phantom v12). The scope of the camera's observation was set in such a way that it covered both the space in front of and behind the multilayer shield (Fig. 2). Due to such a configuration of the camera, it was possible to obtain experimental data, which later served to calculate the speed of the bullet before hitting and after puncture. For the measurement of ballistic resistance, the value of the energy absorbed by the shield during its penetration is accepted, in short *absorbing energy* E_a – the difference of kinetic energies of the bullet before hitting and after puncture the shield specimen. The energy of puncture (absorbing energy) was referred to the thickness of the pack and the basis weight of the tested materials. Moreover, ballistic resistance of the selected specimens we with use a Parabellum bullet were conducted.

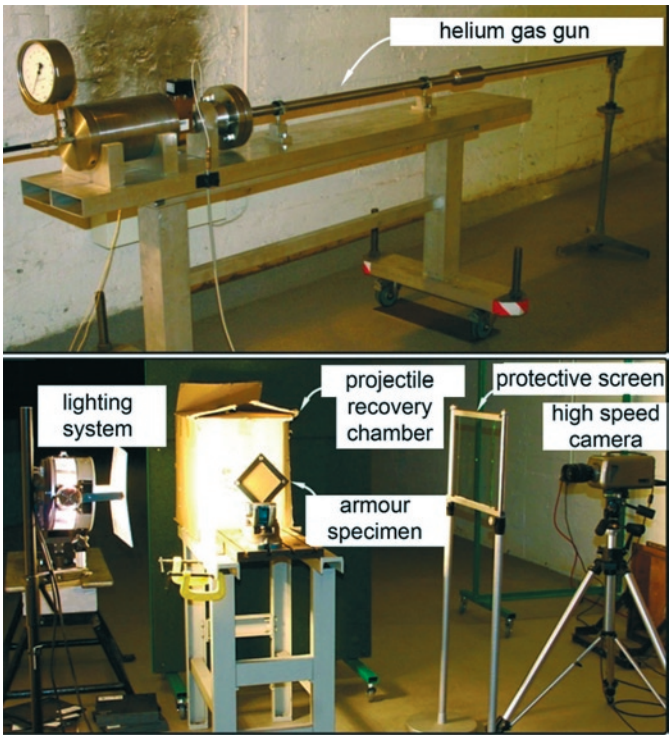


Fig.1. The experimental stand for testing ballistic perforation resistance

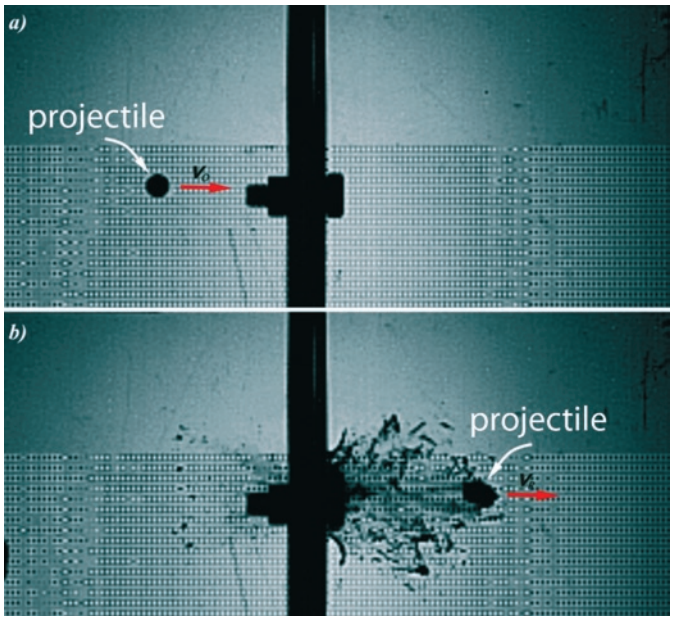


Fig. 2. Observation area of a high speed camera: view of a spherical projectile and an armour specimen before (a) and after (b) impact

2.2. Results and discussion

The composite panels were prepared as aramid-epoxy composites, where the seven layers of the aramid fibre were bonded with the Epidian 57/Z1 adhesive. On the basis of the prepared composite panels, specimens were prepared for ballistic testing. In addition, the specimen was made, consisting of 7 layers of XPS102 fabric, however the fabric layers were not adhesive bonding but stitched together on the edges. The test results are specified in Table 1.

Among the three tested fabrics, the XPS102 fabric is characterized with the best protective properties. The absorbing energy of loose XPS102 fabrics (stitch together on the edge of the specimen) almost doubled the absorbing energy of the laminate, made with this fabric.

The research also is conducted into FML (Fibre Metal Laminates), consisting of alternate layers of thin metal sheets (8 layers) and aramid fabrics (7 layers), adhesive bonding with Epidian 57/Z1. The test results are specified in Table 2.

The absorbing energy, in relation to the basis weight of the tested FML, in each case, proved to be lower than the absorbing energy of the laminates.

Another stage of the research was connected with testing FML materials, which had one layer of SiC ceramics adhesive bonded, shielded with one layer of carbon fibre, oversaturated with Epidian 57/Z1 adhesive. In this case, the penetrators fired from the helium canon (steel bullets) did not break through the first tested specimens (Table 3). Therefore, the two remaining specimens underwent ballistic testing by Parabellum munitions, whose energy of 490 J, exceeds the energy of steel bullets, hitting the examined materials. As expected, the Parabellum bullets did not pierce the tested composites (Fig. 3).

Table 1. Ballisic resistance of laminates (L) and loose aramid layers (7L)

Material	Thickness [mm]	Density [g/cm3]	Absorbing energy [J]	Energy/thickness [J/mm]	Basis weight [kg/m^2]	Energy/ basis weight
L XPS102	3.90	1.18	80.49	20.64	4.60	17.49
L Microflex	2.15	1.23	39.94	18.58	2.64	15.11
L T750	4.10	1.05	72.71	17.73	4.30	16.89
XPS102-7L	3.2 0	1.18	118.18	36.93	3.78	31.30

Table 2. Ballistic resistance of FML composites on the basis of 2024T3 alloy and aramid fibres

Material	Thickness [mm]	Density [g/cm^3]	Absorbing energy [J]	Energy/thickness [J/mm]	Basis weight [kg/m^2]	Energy/ basis weight
FML CT 709	4.10	2.03	100.37	24.48	8.32	12.06
FML XPS102	6.35	1.69	159.33	26.67	10.73	15.78
FML Microflex	4.65	1.91	112.32	24.15	8.89	12.64
FML T 750	6.79	1.69	182.17	26.83	11.48	15.88

Table 3. Ballistic resistance of FML composites, on the basis of 2024T3 alloy and aramid fabrics with Si_2C type ceramics

Material	Basis weight with ceramics [g/cm^2]	Thickness with ceramics [mm]	Thickness [g/cm^3]	Speed [m/s]	Absorbing energy [J]	Energy/ thickness [J/mm]	Energy/ basis weight
FML-C CT709	22.80	8.75	2.61	655	>448	>51.0	> 19.54
FML-C XPS102	24.55	11.00	2.23	657	>451	>41.0	> 18.39
FML-C Microflex	23.36	9.30	2.51	lead bullet Parabellum ~ 490 J		> 52.7	> 20.98
FML-C T750	25.96	11.44	2.27	lead bullet Parabellum ~ 490 J		> 42.8	> 18.88

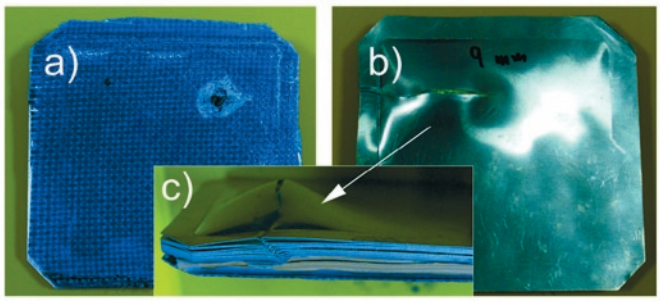


Fig. 3. View of FML-C Microflex specimen after a piercing attempt by a Parabellum bullet: a – view from the ceramics side, b – deformation and cracking of the last metal layer, c – delamination of material

The final stage of the ballistic investigation concerned aramid fabrics, on one side covered by ceramic plates. In the first case, these were seven stitched together layers of CT709 fabric, covered with one layer of Al_2O_3 ceramic plates, adhesive bonded between two sheets made with 2024T3 alloy, 0.3 mm thick. In the second case, a polymer composite on the basis of seven layers of T750 fabric and L285 resin was made, which had one layer of SiC ceramic plates adhesive bonded. The ballistic resistance of shields was checked, prepared in such a way, to piercing by a Parabellum bullet. In both cases, the bullet did not shoot through the specimens (Fig. 4 and 5). The ceramic plates, adhesive bonded directly to the composite made with T750 fabric, with Epidian 57/Z1 adhesive, became separated from the plastically deformed material (Fig. 5). In the case of the plates bonded in between

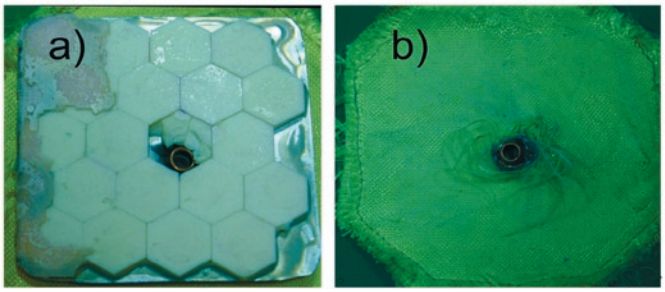


Fig. 4. View of specimen composed of seven stich layers of CT709 fabric and ceramics adhesive bonded between two thin sheets made with AW 2024T3 alloy, after an ballistic attempt with use a Parabellum bullet: a. – view of damaged ceramics b. – view of fabric with a stuck bullet

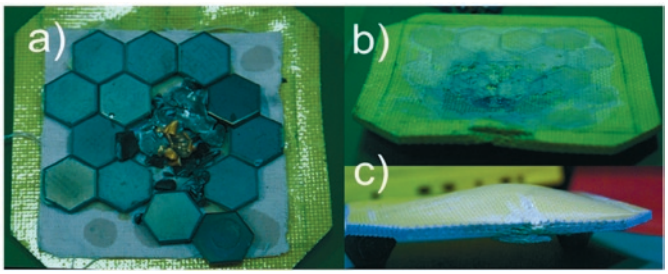


Fig. 5. View of specimen made with T750 fabric composite with a adhesive bonded layer of ceramics, after a ballistic attempt with a Parabellum bullet: a. – view from the side of ceramics b. – view of composite with marks of the separated ceramics c. – permanent deformation of the composite

two layers of thin aluminium alloy sheets, only one ceramic plate was damaged, whereas the remaining ones, which were still bonded to the sheet, shielded the composite (Fig. 4).

Since the tests proved that the ceramic layers considerably increase ballistic of shields, their use in such shields seems to be necessary. Therefore, it was conducted research which assessed the possibility of adhesive bonding a ceramic layer to antiballistic fabrics.

2.3. Adhesive bonding of ceramics to fabrics - materials and method of research

In the research two types of polyethylene fabrics are used, marked as HB50 ROLL 401184C and HB26 ROLL 401400B, four types of aramid fabrics marked as Microflex CT 709, T 750 and XP S102 as well as Al_2O_3 and SiC ceramics.

Taking into account the fact that polyethylene, as thermoplastic component, is characterised with limited adhesion to standard adhesives, the investigation centered around adhesives designed to bond polymer components, including acrylic adhesives: LOCTITE 3038, Scotch-Weld 8010 as well as cyanoacrylate adhesive: BONDICX 01 CX-80, epoxy adhesive: POXIPOL and also silicone adhesives: Professional Black Silicone and Terostat M9399.

Due to lack of norms which refer to specifying the strength of adhesive bonds in materials, which are being investigated, four types of specimens are selected and tested their suitability.

1. The ceramic plate had straps of the selected fabrics bonded on both sides, the straps being 20 mm width and 70 mm long. The fabric endings were fixed in the grips of the strength testing machine and the specimen underwent tension. In the test, the distance between the grips was 70 mm. The adhesive joints of such specimens were mainly loaded to peel off (Fig. 6).

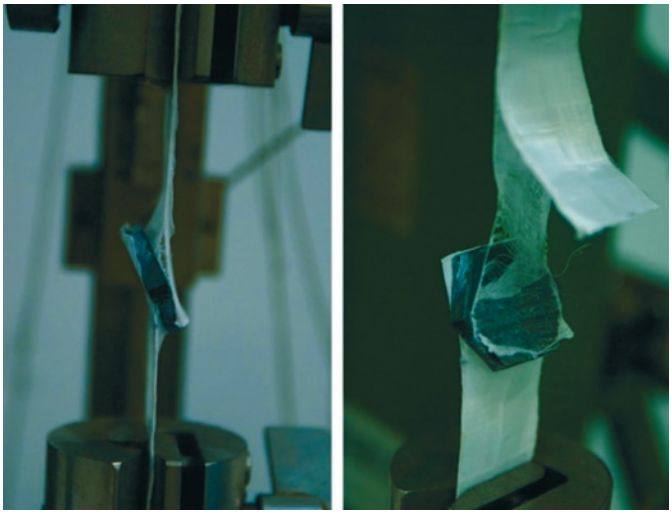


Fig. 6. Manner of fixing specimens no 1

Fig.7. Cylindrical (brass) pieces of the specimen (butt joint) for loading in the tension test.

2. Cylindrical metal elements, front-bonded through a layer of the tested fabric (butt joints) and loaded in the tension test. (Fig.7).
3. The steel cylindrical pieces had ceramic plates bonded with epoxy adhesive Epidian 57/Z1 (Fig. 8). Next the pieces were bonded to the front through a layer of the fabric. In this way butt joints specimens for the testing were prepared and they were next loaded in the tension test.

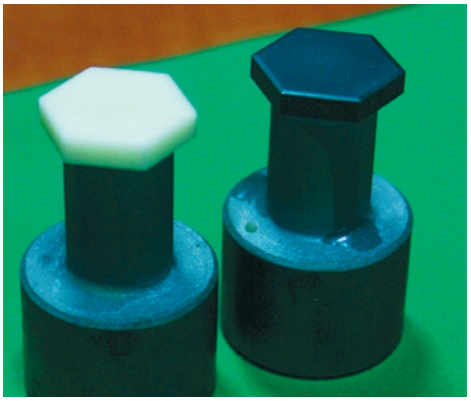


Fig. 8. Steel pieces with bonded ceramic plates

4. In the front sockets of the cylindrical steel specimens had ceramics adhesive bonded (Fig. 9) with Epidian 57/Z1 adhesive. The threaded hole made in the specimen enabled to fix it in the grip of the strength testing machine through a screwed in rod with a tapped thread. The ceramics had fabric straps adhesive bonded, measuring 30 x 150 mm.

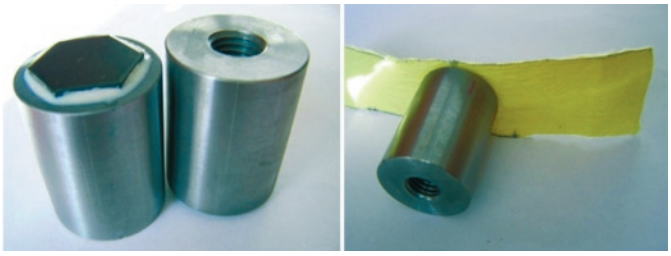


Fig. 9. Specimens with bonded of ceramics and adhesive bonded fabric strap

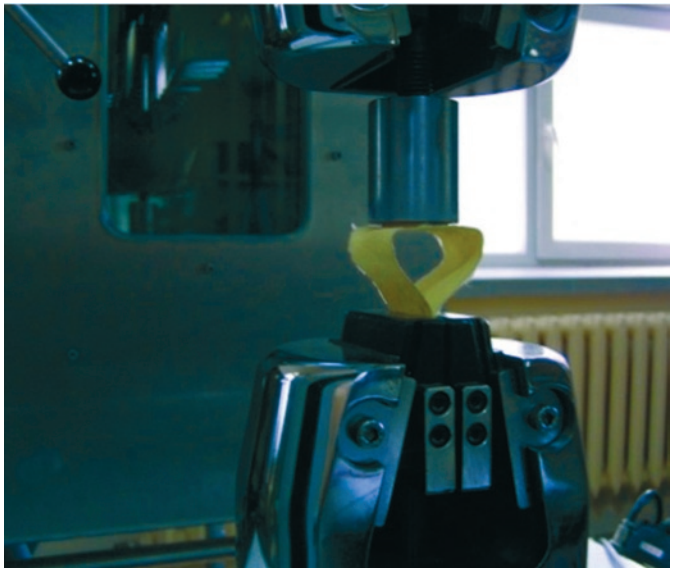


Fig. 10. Specimen fitted in the grips of the strength testing machine

Table 4. The test results of average breaking load of adhesive bonding of polyethylene fabrics to ceramics by means of Bondicx 01 CX80

Bondicx 01 CX80 adhesive		
Fabric HB50 ROLL 401184C		Fabric HB26 ROLL 401400B
SiC	Al_2O_3.	Al_2O_3.
Average breaking load [N]		
19.2±1.8	29±4.4	61.4±5.9

Table 5. The test results of average breaking load of adhesive bonding of polyethylene fabrics to ceramics by means of Poxipol adhesive

Poxipol adhesive	
Fabric HB50 ROLL 401184C	Fabric HB26 ROLL 401400B
Al_2O_3.	
Average breaking load [N]	
22.6±4.3	47.1±3.5

The specimens were fixed in the grips of the strength testing machine (Fig. 10) and loaded at the speed of 10 mm/min. This way of loading caused peeling the fabric off the ceramics.

In the process of adhesive bonding, the ceramic plates were rinsed with isopropyl alcohol and wiped with a cotton cloth until there were no marks on the clean material. The Al_2O_3 ceramics, which was covered with a pink film, needed more thorough rinsing. Due to the hardness of the used ceramics, it was not possible to make the bonded surface rough. The surfaces of the fabrics before bonding were also wiped with a cotton cloth soaked in alcohol. The strength of the specimens of Type 1 (1) was tested on the Louis Shopper strength testing machine, designed for testing fabrics. The strength of the remaining specimens was examined in the Hung Ta HT-2402 strength testing machine.

In the event of a sufficient repetition of the experiment results, they were calculated statistically by determining the confidence interval for the significance level of $1-\alpha = 0.95$.

2.4. Results and discussion

2.4.1. Specimens no 1

In all the cases, the joints were destroyed through separating the thin layer of the fabric (delamination and peeling off the thin layer which is adjacent to the ceramics) – Fig. 6. It means that these fabrics are not fit for adhesive bonding since their tensile strength, perpendicularly to the alignment of the fibres, is low. Moreover, these fabrics cannot be oversaturated even with adhesives of increased fluidity, such as Bondicx 01 CX80. Higher strength of specimens prepared from the HB26 ROLL 401400B fabric resulted exclusively from increased interlayer strength of the fabric itself, and not its better adhesive properties.

2.4.2. Specimens no 2

Polyethylene square-shaped fabrics, with length sides equalling 20 mm, were placed in between the front surfaces of the specimens, which were 16 mm in diameter. It was impossible to measure the strength of the specimens bonded with Bondicx 01 CX80, since they were destroyed through the fabric layer while trying to fix them in the strength testing machine, with forces of 3.5 N – Fig. 11.

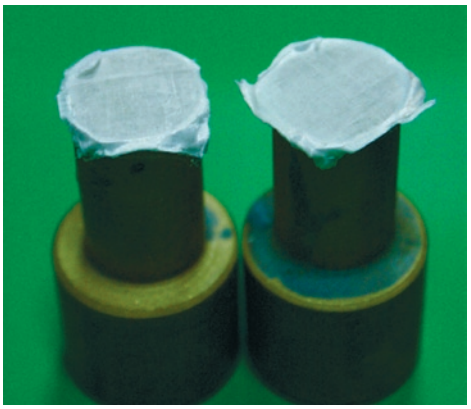


Fig. 11. An example of fabric delamination during an attempt to fix the specimen in the grips of the strength testing machine

Therefore, it was decided to resign from bonded polyethylene fabrics and in the tested butt joints polyethylene foil spacers was used, 0.2 mm thick. Before bonding, the foil was washed with alcohol. The test results are presented in Table 6.

On the basis of the investigation, it was discovered that polyethylene as a thermoplastic component is characterised with low adhesive properties (for instance using a cotton cloth instead of polyethylene, in the test, caused that the strength of the Bondicx 01 adhesive based specimens was increased by over 2.5 times). Thus, the adhesive joints of the polyethylene demonstrate low strength, and the destruction of such joints is typically adhesive in its character. This type of damage occurred in all the cases of the butt joint specimens with polyethylene spacers. Among the tested adhesives, Scotch-Weld 8010 proved to be the most suitable for bonding polyethylene.

Table 6. The test results of average breaking load of butt joints for various adhesives

Adhesives						
Bondicx 01	Poxipol	Loctite3038	Scotch-Weld 8010	Black silicon	Terostat M9399	Bondicx 01 (cotton linen)
Average breaking load [N]						
657	945±49	830±153	1190±184	124±21	208±17	1795±91
Stresses [MPa]						
3.285	4.725 ±0.245	4.15 ±0.765	5.95 ±0.92	0.62 ±0.105	1.04 ±0.085	8.975 ±0.455

Table 7. The test results of average breaking load of butt joints for various adhesives and Microflex fabric

Adhesives						
Bondix 01	Poxipol	Loctite3038	Scotch-Weld 8010	Black silocon	Terostat M9399	Epidian 57/ Z1
Average breaking load [N]						
1684±188	2987±187	0*	232	72.5±12	295±28	2018±185
Stresses [MPa]						
8.38±0.94	14.86±0.93	0*	1.15	0.36±0.06	1.48±0.14	10.04±0.92

** - adhesive not cured within 24 hours*

Table 8. Pull off strength of Bondicx 01 CX80 adhesives in joints with aramid fabrics

BONDICX 01 adhesive				
Fabric	**Microflex**	**T 750**	**CT 709**	**XP S102**
Average breaking load [N]	1684±188	1145±142	1452±90	101.2±6.9
Pull off strength [MPa]	8.38±0.94	5.70±0.71	7.22±0.45	0.503±0.034

Table 9. Pull off strength of Poxipol adhesive in joints with aramid fabrics

Poxipol adhesive			
Fabric	Microflex	T 750	CT 709
Average breaking load [N]	2987±187	1466±140	473.3±49.9
Pull off strength [MPa]	14.86±0.93	7.29±0.70	2.35±0.25

The attempts to bond aramid fibres – XP S102, T750, Microflex and CT709 – were preceded with a preliminary experiment, where small pieces of the fabrics were covered with Epidian 57/Z adhesive in order to check whether these fabrics become oversaturated and whether the saturant demonstrated good adhesion to the fabrics. Out of the tested fabrics, it appeared that only Microflex fabric became thoroughly oversaturated. The T 750 fabric was characterised with the worst adhesive properties, as the curing adhesive practically was not bonded to the fabric. Therefore, the aramid fabric, Microflex type, was selected for the initial strength testing. The results of this phase of the investigation have been specified in Table 7.

The conducted research proved that Loctite 3038 and Scotch Weld 8010 adhesives, designed to bond polyethylene are specialist adhesives and cannot be used for bonding aramid fabrics. The Loctite 3038 adhesive in contact with an aramid fabric and brass did not become cured within 24 hours, however when placed in a polyethylene syringe (remaining only in contact with polyethylene) it became cured at the same time. The investigation proved low pull off strength of silicone sealants – Black Silicone and Terostat M9399 , which results both from poor adhesion to polyethylene and aramids and also their low cohesive strength. A characteristic feature of these adhesive components is their elasticity after curing.

In the further experimental test, the Bondicx 01 CX80 and Poxipol adhesives are used, since they were characterised with sufficiently high pull off strength in the bonding of Microflex fabric. The investigation results of this phase of the investigation have been specified in Tables 8 and 9.

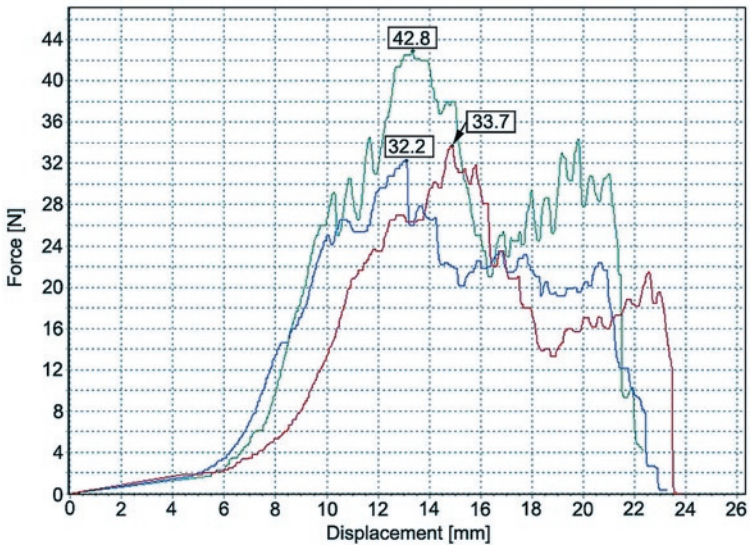


Fig. 12. Examples of three measurements of breaking load to separating Microflex from Al_2O_3 ceramics, bonded with Bondicx 01 adhesive (the value of maximum forces equals 32, 43 and 34 N)

2.4.3. Specimens no 3

The aim of the conducted investigation was to check whether Al_2O_3 and SiC ceramics are characterised with comparable adhesive properties. Specimens no 3 were face-bonded through one layer of Microflex fabric with Bondicx and Poxipol adhesives. The investigation results are specified in Table 10.

Table 10. Comparison of specimens no 3 pull off strength, prepared on the basis of SiC and Al_2O_3 ceramics (Microflex fabric)

Bondicx 01 CX80 adhesive		Poxipol adhesive	
SiC	Al_2O_3	SiC	Al_2O_3
Breaking load [N]			
2522.5*	1994±540	1860±91	2585±548
Stresses [MPa] (A = 346.4 mm^2)			
7.28	5.76±1.56	5.37±0.27	7.46±1.58

**– Destruction of four out of five tested samples occurred on the surface between a steel sample and SiC ceramics, bonded with Epidian 57/Z1. Thus, the strength was measured on the basis of two highest values of the breaking load.*

Table 11. Peel off strength of adhesive joints which bond Microflex fabric to ceramics

Bondicx 01 adhesive		**Poxipol adhesive**	
SiC	**$Al._2O_3$**	**SiC**	**$Al._2O_3$**
Maximum breaking load [N]			
34.75±10.31	34.22±5.38	18.8±1.62	24.2±4.6
Peel off strength [N/mm]			
1,96±0,56	1,95±0,31	1,07±0.09	1,38±0,26

The Bondicx 01 CX80 adhesive demonstrated better adhesion to SiC ceramics, whereas Poxipol adhesive to Al_2O_3 ceramics. It was also proved that Epidian 57/Z1 epoxy adhesive was characterised by better adhesion to Al_2O_3 ceramics.

2.4.4. Specimens no 4

The investigation conducted with specimens no 4 concerned ceramics SiC and Al_2O_3 as well as Microflex fabric and two adhesives: Bondicx 01 and Poxipol. As the adhesive layers of joints were loaded to peel off, the changing forces in the function of displacement the cross beam of strength testing machine was registered (Fig. 12) and the values of maximum forces obtained in the attempts are compared. The investigation results have been listed in Table 11.

The joints where Bondicx 01 adhesive was used, regardless of the type of ceramics, were characterised with heightened strength, as compared to joints formed on the basis of Poxipol adhesive.

3. Conclusions

1. The batches of loose aramid fabrics are characterised with higher ballistic resistance compared to polymer composites made with such fabrics and multilayered composites, built up of thin metal layers and aramid fabrics.

5. Ceramic layers considerably raise ballistic resistance of shields. It seems that their exploitation in such shields is necessary.
6. The tested polyethylene fabrics and the aramid fabric XPS102 cannot be used in adhesive bonding with ceramic plates due to their structure (they consist of thin layers of the material, which is poorly connected), resulting in delamination of the fabrics with low extending loads, perpendicularly to their surface, and inability to oversaturate.
7. The highest pull off strength (14 MPa) was achieved by joining the Microflex fabric with a metal specimen by Poxipol adhesive.
8. It appears that among all the tested specimens, specimens no 4 best reflect loading of ceramic plates, bonded to the fabrics, which are used to manufacture of bulletproof jacket. The research into such specimens indicated that, out of the two examined adhesives, Bondicx 01 adhesive is better suited for such applications than the Poxipol epoxy adhesive.
9. Adhesive bonding of ceramic plates between two thin sheets of aluminium alloy, and not joining them directly with aramid fabrics, in the event of a bullet impact, ensures the destruction of a small area of the ceramic layer, thus decreasing the protective properties of the shield to a small extent.

References

1. Cegła M. Habaj W, Podgórzak P. Development of Lightweight Bulletproof Vest Inserts with Increased Protection Capability. Problemy Mechatroniki 2014; 5(17): 23-33.
2. Dekel E, Rosenberg Z. Terminal Ballistics. Berlin Heidelberg: Springer-Verlag, 2012.
3. Formanek B, Jóźwiak S. Szczucka-Lasota B, Dolata-Groszc A, Bojar Z. Intermetallic alloys with oxide particles and technological concept for high loaded materials. Journal of Materials Processing Technology 2005; 162: 46-51, https://doi.org/10.1016/j.jmatprotec.2005.02.015.
4. Grujicic M, Bell W C, Pandurangan B. Design and material selection guidelines and strategies for transparent armour systems. Materials and Design 2012; 34: 808–819, https://doi.org/10.1016/j.matdes.2011.07.007.
5. Grujicic M, Pandurangan B, D'entremont B. The role of adhesive in the ballistic/structural performance of ceramic/polymer–matrix composite hybrid armor. Materials and Design 2012; 41: 380–393, https://doi.org/10.1016/j.matdes.2012.05.023.
6. Hazell P J. Ceramic Armour: Design and Defeat Mechanisms.Canberra:Argos Press, 2006.
7. López-Puente J,Arias A, Zaera R, Navarro C. The effect of the thickness of the adhesive layer on the ballistic limit of ceramic/metal armours. An experimental and numerical study. International Journal of Impact Engineering2005; 32(1-4): 321-336, https://doi.org/10.1016/j.ijimpeng.2005.07.014.
8. Senderski J, Płonka B, Wiśniewski A, Witkowski Z.Multilayer metal-ceramic passive amour for helicopters and special vehicles. Problemy Techniki Uzbrojenia 2011; 40: 57-64.
9. Wiśniewski A. Armours – construction, design and testing.Warsaw: WNT, 2001.
10. Yong M, Iannucci L, Falzon B G. Efficient modelling and optimisation of hybrid multilayered plates subject to ballistic impact. International Journal of Impact Engineering 2010; 37: 605–624, https://doi.org/10.1016/j.ijimpeng.2009.07.004.

Jan GODZIMIRSKI
Marek ROŚKOWICZ
Faculty of Mechatronics and Aerospace
Military University of Technology
ul. S. Kaliskiego 2, 00-908 Warsaw, Poland

Andrzej KOMOREK
Department of Aviation
Polish Air Force Academy
ul. Dywizjonu 303 35, 08-521 Dęblin, Poland

Emails: jan.godzimirski@wat.edu.pl, marek.roskowicz@wat.edu.pl, a.komorek@wsosp.pl

Article citation info:
STĘPIEŃ Z. A study of factors influencing the formation of harmful deposits in the diesel engine injectors. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 331–337, http://dx.doi.org/10.17531/ein.2017.3.3.

Zbigniew STĘPIEŃ

A STUDY OF FACTORS INFLUENCING THE FORMATION OF HARMFUL DEPOSITS IN THE DIESEL ENGINE INJECTORS

BADANIE CZYNNIKÓW WPŁYWAJĄCYCH NA TWORZENIE SZKODLIWYCH OSADÓW WTRYSKIWACZY PALIWA SILNIKÓW Z ZAPŁONEM SAMOCZYNNYM*

The paper describes various types of deposits of injectors used in CI engines with direct and indirect fuel injection. Various factors have been determined which affect the formation and size of deposits. These factors have also been classified and evaluated in terms of their impact on the formation of deposits precursors and further changes of deposits during the engine life. The results of deposits formation tests have been presented. The external and internal deposits in the High Pressure Common Rail fuel injection systems have been photographed and described. The impact of formed deposits on the diagnostic parameters of injectors has been discussed broadly as it determines the range of injectors' dysfunction and their fit for further use.

Keywords: *compression ignition engines, fuel injectors, external injector deposits, internal injector deposits, hazards caused by deposits.*

W artykule opisano różne typy osadów wtryskiwaczy silników ZS z pośrednim i bezpośrednim wtryskiem paliwa. Określono różnorodne czynniki mające wpływ na tworzenie i zmiany wielkości osadów. Dokonano też klasyfikacji i oceny znaczenia przedmiotowych czynników na powstawanie prekursorów osadów oraz dalsze zmiany osadów w czasie eksploatacji silnika. W zakresie opisu przeprowadzonych badań przedstawiono wyniki procesów tworzenia osadów w warunkach testów silnikowych. Utworzone w ten sposób osady zewnętrzne i wewnętrzne wtryskiwaczy układu wysokociśnieniowego wtrysku paliwa typu common rail były następnie udokumentowane fotograficznie, oraz opisane. Szeroko omówiono też wyniki oceny wpływu wytworzonych osadów wtryskiwaczy na zmiany ich parametrów diagnostycznych, określających zakres dysfunkcji wtryskiwaczy i decydujących o możliwości ich dalszego stosowania.

Słowa kluczowe: *silniki z zapłonem samoczynnym, wtryskiwacze paliwa, osady zewnętrzne wtryskiwaczy, osady wewnętrzne wtryskiwaczy, zagrożenia powodowane osadami.*

1. Preface

The worldwide studies on the formation mechanisms and prevention of deposits in fuel injectors in CI (compression ignition) and SI (spark ignition) engines have been conducted for almost twenty years. The hazards caused by the formation of external (coke) deposits on pintle injectors in CI engines with indirect injection were recognized already in 1990s. The fouling with coke deposits was found on the fuel outflow channel and also on the area around the edge of fuel outlet hole and the cylindrical part of pintle nozzle mating with the fuel injection hole. In case of this type of injectors and indirect injection CI engines, the size and speed of coke deposit formation on injector atomizers is more influenced by the type (design) of the injector itself than by the fuel. The formation process of organic coke deposits is a result of thermal decomposition of fuel during combustion, and the size of deposits can be controlled to some degree by adding detergent-dispersant additives to the fuel. These issues were broadly described by, inter alia, Barker, Snape and Scurr [2], Quigeley, Barbour, Arters and Bush [11], Lacey, Gail, Kientz, Milanovic and Gris [9], and Żak, Ziemiański, Stępień and Wojtasik [19]. On the other hand, in the HPCR (High Pressure Common Rail) injectors, the coke deposits form around and on the edges of fuel outlet holes, but also on the walls inside the fuel channels of injector atomizers. These are usually a mixture of organic and inorganic deposits. Taking into account the fact that the inside diameter of a fuel channel is less than 0.1 mm, the deposits distort the stream of atomized fuel and change its range with an adverse effect on atomization and mixing of fuel with air in combustion chambers as reported by: Birgel et al. [3, 4], Caprotti et al. [5, 6], Quigeley et al. [10] and Stępień [14]. Uncontrolled changes of excess air number in the fuel-air mixture combined with an insufficient atomization of the fuel streams and mixing with exhaust gases at incorrect proportions by the EGR system leads to incomplete combustion. The result is lower engine performance, uneven running, increased emissions of exhaust gases and fuel consumption. The longer range of streams can cause washing the combustion chamber walls with liquid fuel, and this as a result of incomplete fuel evaporation in the air charge leads to incomplete combustion and increased emissions of pollutants, particularly HC. The results of studies conducted by Stępień [15] showed that in the extreme case the deposits formed inside the fuel channels can totally obstruct the hole, particularly when the engine is often stopped and cooled during the operation which allows the deposits to stabilize – Fig. 1.

A very high pressure in the HPCR injection systems, reaching even 300 MPa, allows a very good fuel atomization in the stream injected to the combustion chamber but also causes a significant fuel temperature increase during a rapid depressurization after the fuel squeezes (seeps) through leaks between the injector's mating parts. This, in combination with contaminants which make their way into to the fuel during its production, distribution and storage (particularly metallic elements with catalytic properties), chemical reactions taking place between some fuel additives and biocomponents (FAME) which

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

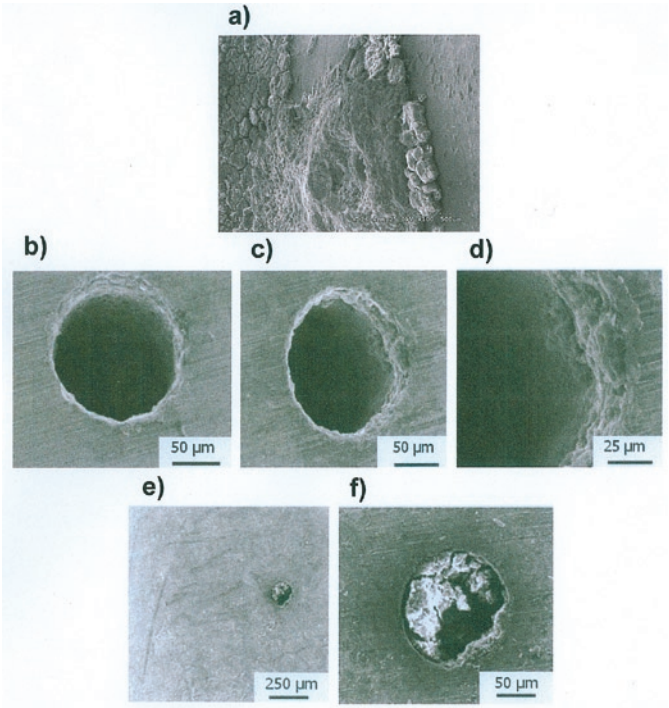


Fig. 1. External coke deposits on an HPCR injectors a) coke deposits formed around the fuel outlet hole b, c, d) coke deposits formed at the hole edge and inside the channel e, f) coke deposits obstructing the hole

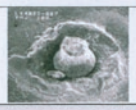
contribute to a reduction of the fuel oxidation stability, leads to the formation of the Internal Diesel Injector Deposits (IDID).

These deposits are divided into a few types and are classified according to various criteria, most frequently by composition determined by means of the FT-IR (Fourier Transform Infrared Spectroscopy) – Table 1. Sometimes the IDID are analysed using other techniques, e.g. SEM (Scanning Electron Microscopy) and EDS (Energy Dispersive X-Ray Spectroscopy), and based on the results they are divided into types (see Yamada et al. [18], Tanaka et al. [16], and Fang and McCormick [7]). That is why various publications use different names for the same deposit types and the number of IDID types differs between authors. Nonetheless, these are always the same deposits from the point of view of composition and formation mechanisms [2, 14, 16]. This has been a motivation for an attempt to systemize the deposit types and factors which influence their formation.

Table 1 presents various types of deposits linked with the factors which influence their formation and size. The table also includes an evaluation of the factors in terms of their impact on the formation of deposit precursors and further changes of deposits during the engine life. Table 1 indicates a large number of varying factors that may affect to a greater or lesser degree the formation of deposits and the size change over time. Some of these factors have an opposing impact on various deposits in different fuel injection systems. For instance, the PIBSI (Polyisobutylene Succinimide) detergent additives widely used in the 1990s prevent the formation of internal coke deposits in pintle injectors in the indirect injection engines. At the same time, such additives are not always effective in preventing the formation of external deposits in modern HPCR systems, and in combination with lubrication additives can cause the formation of polymer amide deposits of the IDID type, particularly in case of PIBSI with small molecular mass, as indicated by Reid and Barker [12].

One of the most important factors causing the formation of deposits in fuel injectors is the fuel oxidation. The content of FAME in fuel has a significant impact on reduction of the fuel oxidation stability. As a result, the oxidation stability of pure FAME is different than of its mixture with diesel fuel at different proportions. In case of engines with indirect fuel injection, where the injector operating temperature is much lower than in the HPCR systems in the HSDI (High Speed Direct Injection) engines, FAME, being a polar solvent, can wash out coke deposits from injectors, particularly at low engine loads [11]. FAME, however, behaves quite differently in modern engines with the HPCR injection systems, with high operating temperatures and

Table 1. Classification of factors influencing the formation of deposits in CI engine injectors

Deposit type and place of formation			External injector deposits		Internal injector desists			
			Indirect injection engines	Direct injection engines	Direct injection engines			
			Coke deposits (organic)	Coke deposits (organic and inorganic)	Carboxylates (with large molecular mass)	Amides	Esters (FAME type)	Esters (CFI type – Cold Flow Improver)
								
Factors influencing the formation of deposits			Impact on the formation and size of deposits (xxx – large / xx – medium / x – small / 0 – insignificant / - - not applicable)					
Fuel	Composition of base fuel	Content of O2, N2 and S	XX	XXX	XX	XX	X	X
	Additives	Detergent (PIBSI-based)	XX	XXX	0	XXX	0	0
		Lubrication	0	X	XXX	XXX	XX	X
		Corrosion inhibitors	0	X	XXX	0	X	X
		Increasing the cetane number	X	X	0	0	XX	X
		CFI	0	0	0	0	0	XXX
	Contaminants	Zn, Cu, Na, K, Ca, Cl, Mg	XX	XXX	XXX	X	0	0
	Biocomponents	FAME	XX	XX	0	0	XXX	0
Engine and fuel injection system	Fuel injection	Injector design	XXX	X	XX	XX	XX	XX
		Operation parameters (HPCR)	-	XX	X	XX	XX	XX
	Systems	EGR, Blow-by	X	XX	0	0	0	0
Operating conditions	Engine load	Low	XX	X	XX	X	X	X
		High	X	XX	X	XX	XX	XX
		Dynamic changes	X	X	X	X	X	X
		Long stoppages	X	XXX	XX	XX	XX	XX

pressures, where fuel heated up to 60 ºC returns many times from the injection systems overflows to the fuel tank, heating up the whole fuel system. Such conditions significantly accelerate the FAME degradation process by oxidation of unsaturated FAME bonds to the peroxide structures which transform into acids and aldehydes. The peroxide structures can also react with other FAME molecules, forming dimers. This can lead to the formation of oligomers which are precursors of the deposits formation if their molecular mass and polarity are sufficiently high. In addition, Fang and McCormick in their studies discovered that the largest deposits are formed in case of mixtures containing from 20% (v/v) to 30% (v/v) of FAME in the diesel fuel [7]. The reason is that the deposits formed from oligomers are highly polar and better soluble in polar FAME, and poorly soluble in hydrocarbons. When the FAME content exceeds 50% (v/v) the fuel becomes sufficiently polar to dissolve deposits formed in the injection system. When the FAME concertation in the fuel is too low (< 5% (v/v)), there are not enough oligomers to form the deposits. The fuel containing from 20% (v/v) to 30% (v/v) of FAME is not sufficiently polar and the oligomers precipitate, forming deposits [7, 11].

The studies conducted to date clearly indicate that trace content of some elements, mostly metallic, accelerate the formation of external and internal deposits (IDID) in injectors, and Zn and Cu to a greater degree intensify the formation of external deposits, and Na, Cl, K, Ca and Mg the formation of IDID, as indicated by the results of studies by Iida [8], Lacey et al. [9] and Shiotani and Goto [13].

Forces between the surface material and the deposit (adhesion) play an important part in keeping the deposits on the surface where they have formed. This applies particularly to the IDID which form on the injector parts that make sliding movements relative to each other. The deposits in these places slow down the injector operation, and consequently can change not only the amount of fuel dosed in a single injection, but also in an uncontrolled way they can alter the time of start and end and duration of individual phases in multiple injection. The result is that the amount of fuel injected to individual cylinders is not equal, and the quality of fuel-air mixture in combustion chambers is not uniform. In case of carboxylate, amide and CFI (Cold Flow Improver) polymer deposits, the adhesion of deposits to the surface is a result of intramolecular forces, and in case of the deposits related to the FAME degradation the chemical bonds come into play as described by Yamada [19].

The formation of deposits is also affected by the engine operating conditions which translate into the engine operation parameters in the test cycle on the test stand. Such conditions are closely related to the fuel composition and the type of deposit, and thus vary depending on the type of deposits we wish to form in the lab and study.

The operational problems relating to the formation of external and internal deposits in the fuel injectors of CI engines were a motivation to perform the study which results are presented below. The aim was to form the external (coke) deposits and internal organic (amide-type) deposits typical for European fuels, and then to evaluate their impact on the injector operation. The innovative character of our study lies in designing a simulation engine test which, in combination with suitably prepared fuels, allowed forming various types of injector deposits on the engine test stand in a short time. The fuel preparation was guided both by our own earlier experiments [14, 16, 18], and by the composition of deposits determined by other research institutes during the actual operation of vehicles [12, 17, 19]. After the formation of deposits in the HPCR injectors, we evaluated, among other things, the diagnostic parameters of the injectors on a professional test stand in order to determine the impact of these deposits on the operational parameters of fuel injectors.

2. Materials used in the tests

In order to accelerate the formation of injector deposits (especially the IDID), the engine was running on the naturally aged, commercially available diesel fuel containing 4.78 % (v/v) of FAME – Table 2. Moreover, added to the fuel as 1-litre premixes were mixtures of chemical compounds which according to the available results of tests performed at various laboratories have the greatest impact on the formation of various IDID [1, 2, 6, 9, 10, 11, 12, 14, 16]. The said

Table 2. Selected physical and chemical properties of fuels used during engine tests

Tested parameter	Test method	Unit	Diesel fuel	B100
Oxidation stability	PN-EN 15751:2014	h	19.8	4.9
Resistance to oxidation	PN-EN ISO 12205:2011+Ap1:2011	g/m^3	5	32
Oxidation stability	ASTM D 7545-14	minute	39	23
Corrosion properties to steel: - temp. 38 °C, test time 5 h - temp. 50 °C, test time 3 h	ASTM D 665-14 Procedure A (distilled water)	-	brak -	- brak
Contaminants	PN-EN 12662:2014 PN-EN 12662:2009	mg/kg	6.0 -	- 182.5
Kinematic viscosity at 40 °C	PN-EN ISO 3104:2004	mm^2/s	3.113	4.540
Distribution of contaminants: - 4, 6 and 14 µm - 2 and 5 µm	ISO 11500:2008 and PN-ISO 4406:2005	-	18/16/14 18/16	23/23/23 23/23
Fractional composition: - up to temp. 250 °C distils - up to temp. 350 °C distils - 95 %(*V/V*) distils up to temp.	PN-EN ISO 3405:2012	%(*V/V*) °C °C	29.6 92.6 358.8	-
Density at 15 °C	PN-EN ISO 12185:2002	kg/m^3	838.8	-
Sulfur content	PN-EN ISO 20846:2012	mg/kg	7.7	-
Water content	PN-EN ISO 12937:2005	mg/kg	65	-
Flash point	PN-EN ISO 2719:2007	°C	69	-
Cetane number	PN-EN 5165:2014	-	52.8	-
Cetane index	PN-EN ISO 4264:2010+A1:2013-07	-	54.1	-
Cold filter plugging point (CFPP)	PN-EN 116:2015	°C	-6	-
Carbon residue from 10% distillation residue	PN-EN ISO 10370:2014	%*(m/m)*	0.048	-
Ash residue	PN-EN ISO 6245:2008	%*(m/m)*	<0.001	-

chemical compounds comprised the ingredients present in detergent and lubrication additives, corrosion inhibitors, additives increasing the cetane number and the contaminants which make their way into to the fuel. As a result, two fuels described below were prepared for the simulation engine tests.

In test No. 1, the propensity for the formation of external and internal deposits was tested on the aged, commercially available diesel fuel (physical and chemical properties are given in Table 2) with addition of 120ppm of PIBSI (Polyisobutylene Succinimide) + 100ppm of hydrogenated dimer of oleic acid + 1ppm of formic acid. This aimed at accelerating the formation of amide deposits by adding PIBSI on which the detergent additives to diesel fuel are based and by adding acidic substances which are ingredients of lubrication additives.

In test No. 2, similarly to test No. 1, the propensity for the formation of external and internal deposits was tested on the aged, commercially available diesel fuel (Table 2) in which the content of biocomponents (4.78% (v/v) FAME) was increased to 10% (v/v) by adding aged B100 (properties given in Table 2). Moreover, 500 ppm of 2 ethyl-hexyl nitrate (2-EHN) + 120 ppm of PIBSI (small molecular mass) were added to the fuel. This aimed at accelerating the formation of amide (polymer-type) deposits by adding PIBSI on which the detergent additives to diesel fuel are based and the formation of ester deposits from aged FAME in tandem with 2 ethyl-hexyl nitrate (2-EHN) added to increase the cetane number.

3. Test methodology

The engine simulation tests were performed on a multi-purpose test stand equipped with a modern, widely used HSDI (High Speed Direct Injection) FORD 2.0i 16V Duratorq TDCi diesel engine coupled with an Alpha 160 AF eddy current brake from AVL, with a control module allowing the programming of the engine operation parameters (rpm and load). The basic technical parameters of the FORD 2.0i 16V Duratorq TDCi are given in Table 3.

Table 3. Selected technical parameters of the FORD 2.0i 16V Duratorq TDCi engine

Engine type	four-stroke, compression ignition
Fuel injection type	Direct fuel injection, common rail (Delphi) electronically controlled, cooperating with the Levanta engine control system
Cylinders arrangement	In-line, vertical
No. of cylinders	4
Injection sequence	1-3-4-2
Type of timing gear	DOHC/4 VPC
Cylinder diameter	86.0 mm
Piston stroke	86.0 mm
Displacement	1998 cm^3
Maximum power	130 KM (96 kW) at 3800 rpm
Max. torque	330 Nm at 1800 rpm
Max. instantaneous speed	4800 rpm
Idle speed	750±20 rpm
Compression ratio	18.2
Filling	turbocharged with intercooler and "overboost" function
Valve clearance	Hydraulic adjustment
Capacity of lubrication system with filter	6.0 dm^3
Complies with emission standard	Euro IV

Table 4. Parameters of the 4-phase engine cycle

Phase	Time [s]	Engine speed [rpm]	Engine load [Nm]
1	30	1000	~0
2	150	1800	100
3	60	3000	70
4	60	1500	50

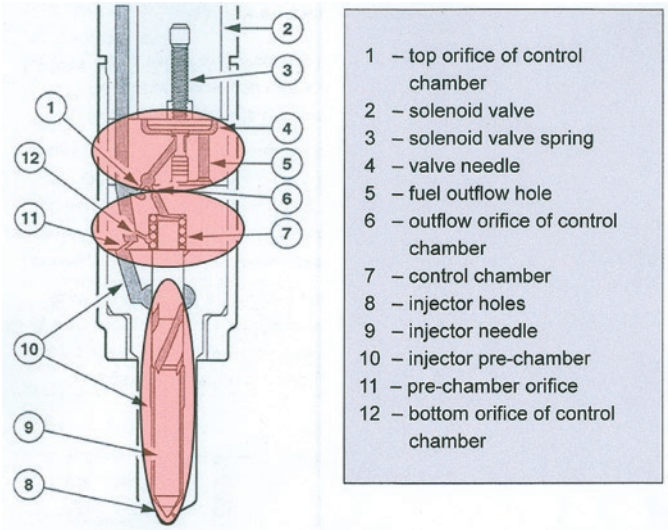


Fig. 2. Longitudinal section of the bottom part of an injector with evaluated components marked

The engine tests were performed in a 4-phase, repeatable cycle which reflected the average engine operating conditions in a low-intensity city traffic. The parameters of the 4-phase cycle are given in Table 4.

The test duration was 100 hours of actual engine operation in 8-hour periods which comprised repetitions of the 4-phase cycle (Table 4) and were separated by 16-hour interruptions during which the formed deposits stabilized.

When the test was complete, the injectors were removed from the engine and two of them were sent for full, professional diagnostics and the other two were left to take photographs of the internal deposits. The necessity for such procedure resulted from the fact that the deposits may be subject to mechanical and chemical damage/changes during the diagnostics and in the final stage they are removed in order to evaluate the parameters and the adjustment possibilities of the "cleaned" injectors. On the other hand, before the diagnostics the injectors must not be dismantled (e.g. for photographing and evaluation of deposits) as this would affect (distort) the diagnostic parameters after the test.

The injector evaluation after the test comprised both the external coke deposits and the IDID. In case of the IDID, the obligatory evaluation was performed on the injector needle, solenoid valve control piston, and control chamber body as components with critical impact on the correct injector operation, and in addition most often subject to damage. The evaluation could have been extended to other components if necessary.

Figure 2 presents components of the HPCR injectors which were evaluated in terms of fouling with the IDID.

The measurements and evaluations of selected diagnostic parameters were performed on the Zapp CRU.2i testing device. The evaluations of each diagnostic parameters are an average value from three measurements.

The test of each injector involved a measurement of a number of different parameters, and then their evaluation in relation to permissible and limit deviations typical for an injector in full good working order, and permissible deviations for an injector in use specified by its manufacturer. Figure 5 presents a list of evaluations of six selected parameters of tested injectors. In addition, each evaluation is presented on a small graph of a parameter of the tested injector in relation to permissible deviations of an injector in full good working order (broken lines) and limit, permissible deviations of an injector in use (full lines). The graphs of parameters of injectors in full good working order were marked in blue, and the parameters outside the permissible limits for injectors in full good working order (but not excluding further use and not requiring repair) were marked in green, and the red colour indicates the injector parameters outside the permissible tolerances for injectors in use. Such injectors with parameters marked in red must be cleaned and/or repaired and then readjusted.

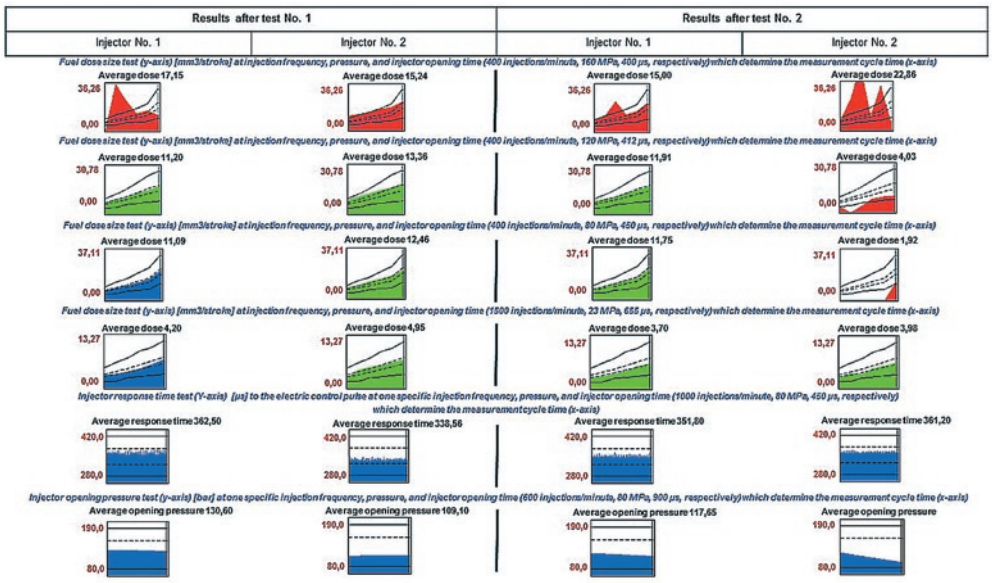


Fig. 5. Diagnostic evaluation results of FORD Duratorq 2.0 TDCi injectors after test No. 1 and 2

4. Test results

Figure 3 presents the tips of injectors after test No.1.

Fig. 3. Photographs of injector tips from FORD Duratorq 2.0 TDCi after test No. 1

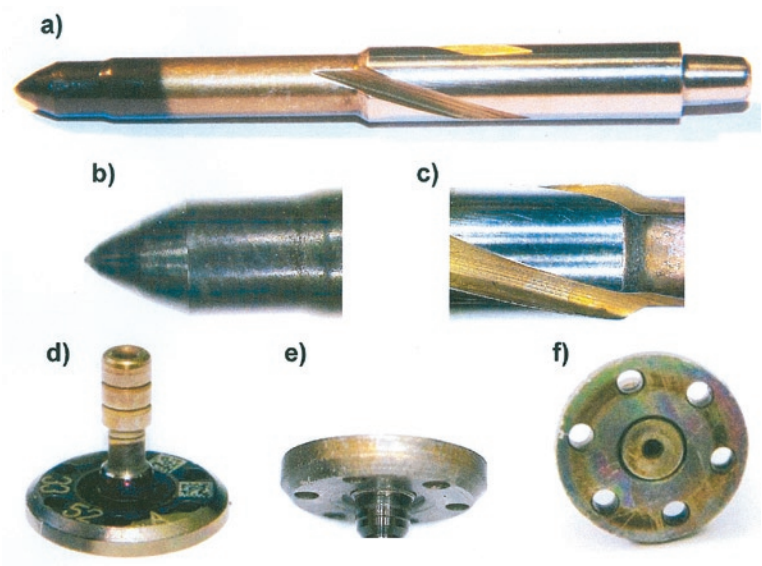


Fig. 4. Images of selected internal parts of the HPCR injectors a) injector needle, b) conical end of injector needle, c) guiding part of injector needle, d, e, f) piston of fuel flow control valve

The injector tips have quite thick layers of uniform, mat, dry, dark brown deposits. The deposits are thicker around the fuel outlet holes. The peeling of deposits is present in some areas – Figure 3. The size and distribution of deposits on the evaluated surfaces is similar. Figure 4 shows selected internal parts of injectors after test No. 1.

Particularly intensive deposits, covering the surfaces in an irregular manner, were found on the conical ends of injector needles and on the cylindrical part of needles directly above the conical part – Fig. 4 a, b. More uniform amide deposits, although of varying thickness, covered the internal surfaces of grooves in the cylindrical guiding part of the needle, and plungers and heads of pistons of the fuel flow control valves – Fig. 4. c, d, e, f.

The next part of the evaluation comprised a measurement of selected diagnostic parameters of two injectors after test No. 1.

The injectors evaluations from both tests presented in Fig. 5 indicate that after test No. 1 the upper limit values of dosed fuel were significantly exceeded at the injection pressure of 160 MPa and the injector opening time of 400 µs. As a result, from the point of view of operational parameters, the injectors were not fit for further use without maintenance-repair operations comprising washing, adjustment and possibly replacement of some internal parts. In addition, there were lesser, permissible deviations in fuel doses at other pressures and injector opening times. This concerned one measuring range in case of injector No. 1 and all measuring ranges in case of injector No. 2 – Fig. 5. The most probable reason of such deterioration of the injectors' diagnostic parameters were internal deposits formed on the working part of the needle which slides on the injector body, and on the plunger of the fuel flow control piston inside the injector. This changed the injector operation time as a result of a delayed or slowed down response of the needle to the electric control pulse, and even caused a possibility of the needle's hanging in the injector body.

Fig. 6 presents the tips of injectors after test No.2.

Mat, dark grey, peeling deposits or varying thickness were formed on the injector tips. The deposits were slightly thicker around the fuel outlet holes – Fig. 6. Figure 7 shows selected internal parts of injectors after test No. 2.

Intensive amide (polymer) deposits accompanied by ester deposits were formed on the injectors inside surfaces. A particularly large amount of deposits was found on conical ends

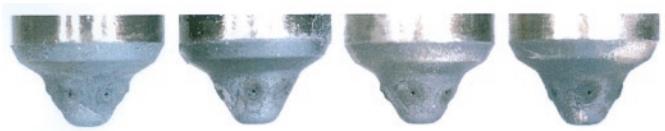


Fig. 6. Photographs of injector tips from FORD Duratorq 2.0 TDCi after test No. 2

of injector needles and on the cylindrical part of needles above the conical part. These deposits had an irregular shape and colour changing from yellow through orange, brown to dark brown. – Fig. 7 b. Intensive polymer deposits were found on the cylindrical parts of the needle washed with fuel, and in grooves in the cylindrical guiding part of the needles, and also in grooves of plungers of the fuel flow control pistons – Fig. 7 c, d, e. The sliding, cylindrical surfaces of needles and plungers feature areas where the deposits were mechanically rubbed during the injector operation – Fig. 7 d, e.

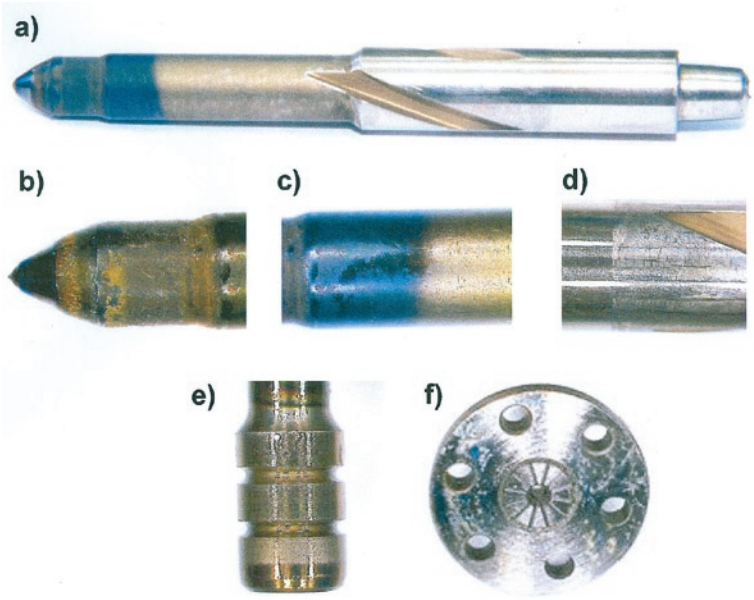


Fig. 7. Images of selected internal parts of the HPCR injectors a) injector needle, b) conical end of injector needle, c) cylindrical part of injector needle washed with fuel, d) guiding part of injector needle, e, f) piston of fuel flow control valve

Graphical results of evaluation of injectors from the FORD Duratorq 2.0 TDCi engine after test No. 2 are presented in Fig. 5. Similarly to test No. 1, in both injectors the upper limit values of dosed fuel were significantly exceeded at the injection pressure of 160 MPa and the injector opening time of 400 µs – Fig. 8.

In case of injector No. 2, the lower limit values of dosed fuel were also exceeded beyond the permitted tolerances at the injection pressure of 120 MPa and the injector opening time of 412 µs, and at the injection pressure of 80 MPa and the injector opening time of 450 µs – Fig. 5. For the same parameters, the deviations in injector No. 1 were lesser, within the tolerance limits. The summary evaluation showed however that both injectors were not fit for further use without maintenance-repair operations comprising washing, adjustment and possibly replacement of some internal parts.

Similarly to test No. 1, the most probable reason of such deterioration of the injectors' diagnostic parameters were internal deposits formed on the working part of the needle which slides on the injector body, and on the plunger of the fuel flow control piston inside the injector. This changed the injector operation time as a result of a delayed or slowed down response of the needle to the electric control pulse, and even caused a possibility of the needle's hanging in the injector body.

5. Conclusions

- The formation of deposits in fuel injectors of the CI engines is influenced by many simultaneously interacting factors. In many cases, specific factors have opposite effects on various groups (types) of deposits, initiating and intensifying the formation of some and at the same time reducing the formation of others.
- The number and complexity of the factors influencing the formation and growth of the IDID in the HPCR injectors of the CI engines still requires more investigation into the their significance and mechanisms of cooperation.
- The IDID formed during the tests caused typical dysfunctions of the HPCR injectors observed during actual operation of vehicles, sometimes making further use impossible.
- The progress in the area of design and technology of the piston combustion engines and the fuel injection systems used in them, as well as the changing fuel production technologies will require the development and use of increasingly more effective detergent additives of multidirectional action. In addition to the design modifications of injectors, such additives are the most effective means to prevent the formation of external and internal injector deposits.
- The evaluation of diagnostic parameters of the HPCR injectors allowed a widened, precise determination of injector dysfunctions caused by the formed deposits.

Abbreviations

CFI Cold Flow Improver
EDS Energy Dispersive X-Ray Spectroscopy
2-EHN 2 Ethyl-Hexyl Nitrate
FAME Fatty Acid Methyl Esters
FT-IR Fourier Transform Infrared Spectroscopy
GDI Gasoline Direct Injection
HPCR High Pressure Common Rail
HSDI High Speed Direct Injection
IDID Internal Diesel Injector Deposit
MPFI Multi Port Fuel Injection
PAH Polynuclear Aromatic Hydrocarbons
EGR Exhaust Gas Recirculation
PIBSI Polyisobutylene Succinimide
SEM Scanning Electron Microscopy
SPI Single Point Injection
PCV Positive Crankcase Ventilation
EDM Electrical Discharge Machining
SI Spark Ignition
CI Compression Ignition

References

1. Baker J, Cook S. Sodium Contamination of Diesel Fuel, its Interaction with Fuel Additives and the Resultant Effects on Filter Plugging and Injector Fouling. 2013; SAE 2013-01-2687.
2. Baker J, Snape C, Scurr D. Diesel Deposits. 2013; 9th International Colloquium Fuels Conventional and Future Energy for Automobiles. 15-17 January 2013.
3. Brigel A, Ladommatos N, Aleiferis P, Zülch S, et al. Deposit Formation in the Holes of Diesel Injector Nozzles: A Critical Review. 2008; SAE Technical Paper No. 2008-01-2383.
4. Birgel A, Ladommatos N, Aleiferis P, Milovanovic N, Lacy P, Richards P. Investigations on Deposit Formation in the Holes of Diesel Injector Nozzles. 2011; SAE 2011-01-1924.
5. Caprotti R, Breakspear A, Graupner O, Klaua T. Diesel Injector Deposits Potential on Future Fuelling Systems. 2006; SAE Technical Paper No. 2006-01-3359.
6. Caprotti R, Bhatti N, Balfour G. Deposit Control in Modern Diesel Fuel Injection Systems. 2010; SAE Technical Paper No 2010-01-2250.
7. Fang H L, McCormick R L. Spectroscopic Study of Biodiesel Degradation Pathways. 2006; SAE Technical Paper No.2006-01-3300.
8. Iida Y. Biodiesel Studies in Japan. 2012; CEN/TC19/WG24, 22 May 2012.
9. Lacey P, Gail S, Kientz J M, Milanovic N, Gris C. Internal Injector Deposits. 2011; SAE Technical Paper No. 2011-01-1925.
10. Quigeley R, Barbour R, Fahey E, Arters D, Wetzel W, Ray J. A Study of The Internal Diesel Injector Deposit Phenomenon. 2009; TAE Fuels 7th Annual Colloquium – 2009.
11. Quigley R, Barbour R, Arters D, Bush J. Understanding the Spectrum of Diesel Injector Deposits. 2013; 9th International Colloquium Fuels Conventional and Future Energy for Automobiles. 15-17 January 2013.
12. Reid J, Barker J. Understanding Polyisobutylene Succinimides (PIBSI) and Internal Diesel Injector Deposits. 2013; SAE Technical Paper No. 2013-01-2682.
13. Shiotani H, Goto S. Studies of Fuel Properties and Oxidation Stability of Biodiesel Fuel. 2007; SAE Technical Paper No. 2007-01-0073.
14. Stępień Z. The reasons and adverse effect of internal diesel injector deposits formation – Przyczyny powstawania i szkodliwy wpływ wewnętrznych osadów we wtryskiwaczach silników o zapłonie samoczynnym. Combustion Engines (Silniki Spalinowe) 2014; 1 (156) : 20 – 29.
15. Stępień Z. Investigations of injector deposits in modern diesel engines - Badanie osadów wtryskiwaczy nowoczesnych silników z zapłonem samoczynnym. Combustion Engines (Silniki Spalinowe), 2016; 2 (165) : 9 – 20.
16. Stępień Z. Ewolucja metodyki oceny zanieczyszczenia rozpylaczy silników o zapłonie samoczynnym. Nafta-Gaz, 2014; 10 : 707 – 716.
17. Tanaka A, Yamada K, Omori T, Bunne S, Hosokawa K. Inner Diesel Injector Deposit Formation Mechanism. 2013; SAE 2013-01-2661
18. Urzędowska W, Stępień Z. Prediction of threats caused by high FAME diesel fuel blend stability for engine injector operation –Fuel Processing Technology 2016; 142: 403-410. https://doi.org/10.1016/j.fuproc.2015.11.001.
19. Yamada K, Bunne S, Omori T. Diesel Injector Deposit. 2015 Powertrains, Fuels and Lubricants Meeting TWS2 2nd, September, 2015.
20. Żak G, Ziemiański L, Stępień Z, Wojtasik M. Problemy związane z tworzeniem się osadów na elementach układów wtryskowych nowoczesnych silników Diesla – przyczyny, metody badań, przeciwdziałanie. Nafta-Gaz 2013; 9: 702 – 708.

Zbigniew STĘPIEŃ
Oil and Gas Institute – National Research Institute
Lubicz 25 A, 31-503 Krakow, Poland
E-mail: stepien@inig.pl

Article citation info:
HAWRYLUK M, MRZYGŁÓD B. A durability analysis of forging tools for different operating conditions with application of a decision support system based on artificial neural networks (ANN). Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 338–348, http://dx.doi.org/10.17531/ein.2017.3.4.

Marek HAWRYLUK
Barbara MRZYGŁÓD

A DURABILITY ANALYSIS OF FORGING TOOLS FOR DIFFERENT OPERATING CONDITIONS WITH APPLICATION OF A DECISION SUPPORT SYSTEM BASED ON ARTIFICIAL NEURAL NETWORKS (ANN)

ANALIZA TRWAŁOŚCI NARZĘDZI KUŹNICZYCH DLA RÓŻNYCH WARUNKÓW EKSPLOATACJI Z WYKORZYSTANIEM SYSTEMU WSPOMAGANIA DECYZJI OPARTEGO O SZTUCZNE SIECI NEURONOWE*

The paper presents the results of research concerning the percentage participation of destructive mechanisms for two typical variants of exploitation of forging tools: lubricated and cooled, and without lubrication. Discussed results come from the developed by the authors the decision support system (SEPEK-2) based on artificial neural network. The knowledge about the durability of forging tools needed for learning artificial neural network was included in the training data set, from comprehensive studies, carried out in industrial conditions. Set of training data set included 450 records of knowledge. The paper presents the process of acquiring knowledge, adopted neural network architecture and parameters developed network. Carried out a global analysis of the results generated by the developed system for the durability of forging tools treated as the maximum number of produced forgings to their destruction (from 0 to 25,000 items), showed that for the lubricated and cooling tools the dominant mechanism is thermo-mechanical fatigue, and do not abrasive wear, which actually dominates in the process of forging tools for uncooled and unlubricated tools. It should be emphasized that the overwhelming majority of studies in this area is attributed that to abrasive wear is dominant, and as shown by the results of research and analysis for the selected representative forging processes, with the use of decision support system based on ANN, the fatigue a thermo-mechanical is dominant in these processes. However, due to the easy measurability and commonly used models wear, based on the model of Archard, it is abrasive wear assigned the largest participation. In fact, for the tool lubricated and cooled tools a thermo-mechanical fatigue intensifies this effect attributed to abrasive wear. While the generally accepted view is correct, in the case of tools unlubricated, as confirmed by the analysis using ANN.

Keywords: *artificial neural network; decision support system; die forging; durability of forging tools; wear and destructive mechanisms.*

W pracy przedstawiono wyniki badań, dotyczące, procentowego udziału mechanizmów niszczących dla dwóch typowych wariantów eksploatacji narzędzi kuźniczych: smarowanych i chłodzonych oraz bez smarowania. Prezentowane wyniki pochodzą z opracowanego przez autorów systemu wspomagania decyzji (SEPEK-2) działającego w oparciu o sztuczną sieć neuronową. Wiedza o analizowanym zagadnieniu trwałości narzędzi kuźniczych, potrzebna do procesu uczenia sztucznej sieci neuronowej zawarta była w zestawie danych uczących, pochodzących z kompleksowych badań, zrealizowanych w warunkach przemysłowych. Zestaw danych uczących obejmował zbiór 450 rekordów wiedzy. W pracy przestawiono proces pozyskiwania wiedzy, przyjętą architekturę sieci neuronowej oraz parametry opracowanej sieci. Przeprowadzona globalna analiza wyników generowanych przez opracowany system, dla trwałości traktowanej jako zwiększająca się liczba odkuwek (od 0 do 25000 sztuk), wykazała że dla narzędzi smarowanych i chłodzonych dominującym mechanizmem jest zmęczenie cieplno-mechaniczne, a nie zużycie ścierne, które rzeczywiście dominuje w procesach kucia dla narzędzi niechłodzonych i niesmarowanych. Należy podkreślić, że zdecydowana większość opracowań z tego obszaru przypisuje, że to zużycie ścierne jest dominujące, a jak wykazały wyniki badań i analiz dla wybranych reprezentatywnych procesów kucia, przy wykorzystaniu systemu wspomagania decyzji opartego o SNN, to zmęczenie cieplno-mechaniczne jest dominujące w tych procesach. Jednakże ze względu na łatwą mierzalność oraz popularnie stosowane modele zużycia ściernego, bazujące na modelu Archarda, to właśnie zużyciu ściernemu przypisuję się największy udział, choć w rzeczywistości dla narzędzi smarowanych i chłodzonych zmęczenie cieplno-mechaniczne wzmaga ów efekt przypisywany zużyciu ściernemu. Natomiast ogólnie przyjęty pogląd jest słuszny, w przypadku narzędzi niesmarowanych. co potwierdziły także analizy przy wykorzystaniu SNN.

Słowa kluczowe: *sztuczna sieć neuronowa; system wspomagania decyzji; kucie matrycowe; trwałość narzędzi kuźniczych; zużycie i mechanizmy niszczące.*

1. Introduction

Forging tools used in semi-hot and hot die forging processes characterize in a relatively low durability, which, in turn, significantly affects the quality and cost of the fabrication of forgings. The low durability of forging tools is mostly caused by the extreme conditions of the industrial hot forging processes, resulting from a simultaneous occurrence of many complex phenomena and degradation mechanisms. It is a difficult and unresolved issue, both scientifically and economically. The degradation and wear of forging instrumentation during its performance constitutes a significant part of the production costs [8, 9]. At present, it is estimated that the costs of tools can constitute up to 8-15 % of the total production costs, and in extreme cases, with small production series, even 30%. In practice, considering the

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

time needed to replace the used instrumentation or in the case of its unexpected degradation, these costs can rise even up to 40%. Also, tool wear significantly lowers the quality of the fabricated forgings. The most common forging defects caused by tool wear are errors in the filling of the die, that is: incomplete forgings, laps, burrs, curvatures, scratches, delaminations, micro- and macro-cracks, etc., which, in turn, affects the functionality of the final product obtained from the forging [11].

In the literature you can find information, that statistically 70% of the forging dies is withdrawn from production due to the loss dimensions – due to abrasive wear and plastic deformation [16], 25% – as a consequence of fatigue cracks and only 5% for other reasons (non-compliance with the technology, construction and material defects, thermal and thermo-chemical treatment defects, etc.) [13]. What is more, the continuous market competition forces the producers of forged goods to constantly lower the costs and produce high quality forgings, and this generates high interest in the problem of tool life improvement [7,12]. More and more often both manufacturers of forging instruments, as well as research centers and research and development use a variety of IT tools and advanced research and develop modern methods allowing both to analyze and to increase the durability of forging tools. Studying the destructive mechanisms and other phenomena accompanying ago and life prediction tools in the forging process is a very difficult and complex issue, and additionally commonly used modes of failure are not described satisfactorily in an analytical way [1, 4, 15, 19, 23].

Predicting the life of a forging tool used in a die forging process and pointing to the dominant wear mechanisms and their effect on the 'life time' of the tool are very important, and they still constitute a not fully resolved problem, in respect of both science and economy [7, 11]. Developing methods which allow for determining the wear (failure) of tools and predicting their life is justified by many factors, such as: continuous perfection of the forging technology, lowering the costs of manufacturing the tools themselves or the cost per unit of the production of one forging, as well as ecological aspects [7, 11, 12]. Modern IT technologies are constantly offering new methods and approaches making it possible to partially replace the costly and time-consuming material experiments with a virtual experiment. Also, new formalisms of representing knowledge in computer systems are being developed, such as: the graph theory, fuzzy logic, artificial neural networks, or genetic algorithms, thus providing the opportunity to construct expert systems supporting various areas of human activity. Expert systems are widely applied mainly as advisory systems for the tasks of identification, classification, control, simulation and diagnostics. [2, 3, 5, 14, 18, 19, 21, 24, 29].

The authors made an attempt at elaborating an original decision support system, whose basic task is to predict the life of the forging tools used in hot die forging processes, as well as to identify the basic failure mechanisms. During the many years of research concerning the analysis of the effect of the particular failure mechanisms on the life of forging tools, certain experimental data have been collected. The data are an excellent source of training data for artificial neural networks. The effects of these studies are presented in this article. Thanks to the developed decision support system based on artificial neural networks (ANN) was achieved interesting research results regarding the real participation of major destructive mechanisms for lubricated and cooled tools as well for unlubricated tools.

2. Expert systems in industrial applications

Modern IT technologies are constantly offering new methods and approaches making it possible to partially replace the costly and time-consuming material experiments with a virtual experiment.

An expert systems are widely used mainly as an advisory systems for the tasks of identification, classification, control, simulation, diagnostics. The broadest range of expert systems applications are in technology, inter alia, to diagnose the state of equipment and failure analysis of experimental data, the design of the control devices, design and technology. The literature often attempts at elaborating expert systems, including ones for the optimization of forging processes. These systems have different aims and take advantage of various formal methods of the system knowledge representation. Katayama et al. developed an expert system to design a cold forging process. Fuzzy logic was used to formulate the principles of the database of this system [14]. Fuzzy logic was also used to develop an expert system for predicting the analysis results with the finite element method when solving the problem of rubber cylinder compression [24]. Gangopadhyay et al. elaborated an expert system for predicting loads and axial stresses during forging [3]. Artificial neural networks have been applied to solve many problems. The application of the finite element method and intelligent system techniques to predict the applied force during the radial forging process was studied in [2]. An artificial neural network was also applied to acquire the relationships between the mechanical properties and the deformation technological parameters of the TC11 titanium alloy, with the use of the data from the isothermal compression test and the conventional tensile test of the forged TC11 titanium alloy at room temperature [18]. In another work [25], the forging and heat treatment experiments of the Ti-6Al-4V alloy were conducted with various experimental parameters, including the forging temperature, deformation degree and heat treatment. On the basis of the obtained experimental data, the optimal model of the hot processing parameters of the Ti-6Al-4V alloy was established with the help of a combination of artificial n neural networks and genetic algorithms.

2.1. Artifical neural networks

Artificial neural networks (ANN) are seen as a separate field of knowledge, offering control and decision-making tools, which perform not only the tasks of approximation, classification and pattern recognition, but also prediction, control and association [20]. ANN are used as advanced analytical tools, useful mainly when it is necessary to model phenomena which are highly non-linear in character, multidimensional functional dependences etc. [29]. Their advantage is the ability to determine the dependences between variables, by way of multiple representations of the training case network or on the basis of the similarities among the cases.

A neural network consists of connections of many neurons arranged in layers. The first layer is the input layer, which does not perform any calculations, and its task is to provide the input signal for the consecutive layers. There are intermediate (so called, hidden) layers. The number of hidden layers depends on the degree of complexity of the modelled function. Most of the existing solutions apply one hidden layer. Within one layer, the neurons are not interconnected. The output layer is the layer which provides the result of the network operation for the analyzed output parameters. The signal transmission in the network usually takes place in the direction from the input to the output, without feedback. There are many types and sorts of neural networks, which differ in the structure and principles of operation [27, 28]. The most popular one is the network architecture connected with the concept of Multi-Layer Perceptron (MLP) [20].

The neural network training methods, broadly discussed in the literature [20, 27, 26], consist in a cyclic update of the network weights based on the information on the objective function gradient and the minimization direction determined in each step. There are many algorithms of MLP network training, such as: the error back propagation algorithm, the steepest descent algorithm, the variable metric algorithm, the Levenberg-Marquardt algorithm, the Gauss-Newton algorithm (the non-linear least square method) and the coupled gradient algorithm. The most effective methods are thought to be the ones

based on the Newton algorithm, which are known as the variable metric methods or quasi-Newton methods. They include: the BFGS (Broyden-Fletcher-Goldfarb-Shanno) method, used in the network elaborated by the authors, and the DFP (Davidon-Fletcher-Powell) method.

Appropriately designed neural networks can formulate dependences occurring between the phenomenon parameters during the learning process. The learning process is an iterative one, repeated multiple times, step by step, whose basic objective is optimization of the network parameters, i.e. the weight coefficients. The objective of neural network training is such selection of its topology and parameters which will assure minimization of error when the output value is set.

3. Methodology section

The article focuses on the analysis of the results, which achieved from the system, developed by the authors, to analyze the durability of forging tools. To build a system used the comprehensive results of long-term experimental research conducted in industrial conditions, relating to three selected representative hot die forging processes. Analysis of the source data allowed, for knowledge representation in the system (as a formal tool) accept artificial neural networks. The ANN is used as an advanced analysis tools useful mainly in cases when it is necessary of phenomena modeling with highly nonlinear nature multidimensional depending function, as is the case in processes analyzed. Another advantage of neural networks is their ability to independently explore the relationships between variables either by presenting multiple network cases learners. Scheme of work and methodology shown in Figure 1.

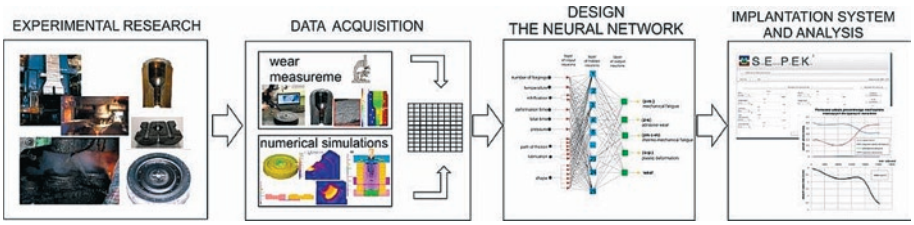


Fig. 1. The flowchart of research methodology

The first version of the system (SEPEK-1), applies fuzzy logic as the form of knowledge representation [5]. This choice was mainly determined by the character of the obtained source data, which are the results of measurements and observations as well as results of computer simulations. They are mainly incomplete data (representing only particular cases for a given number of forgings under given conditions). They are also uncertain as burdened by measurement and observation errors. Fuzzy logic is a formalism which finds its application in the processing of this type of knowledge. The conclusions drawn by the system were verified by experts and recognized as correct and true. The elaborated system, with the use of a fuzzy knowledge base, works with an error at the level of 10%, but the collected data did not fully cover all the analyzed areas. In some cases, the reply generated by the system was that there was no knowledge (rules) based on which a decision could be made.

In the search for forms of knowledge representation which, based on the elaborated experimental data, will make it possible to construct a model characterizing in a lower error and cover all the analyzed areas, Therefore the authors decided to look for other forms of knowledge representation, which on the basis of collected and developed experimental data allow to build a model characterized by even lower error and covering all analyzed areas. Hence it was decided to use artificial neural networks and to develop a new decision support system SEPEK-2.

3.1. Description of experimental data used in model construction

To build a knowledge module, three types of industrial forging processes were selected, so that these processes would occur under a variety of conditions that determine the durability of forging tools. The authors' experience suggests that such differentiation of forging conditions (due to temperature, pressure, path friction, thermo-chemical, tribological conditions, etc.) can be obtained by way of selecting:

I Hot forging of a front wheel in open dies.

II Precision forging of a CVJB (constant velocity joint) housing in closed die.

III Forging of a fastener in a closed die without lubrication and cooling tools.

I Hot forging of a front wheel in open dies – die inserts, heated to 250°C, and cooled and nitrided only in operation 2 and 3, initial billet heated inductively to 1150°C. A wheel forging (Fig 2c), after toothing, is used in the gear box as the front wheel of the reverse gear in motorcars. The analyzed process is realized on a crank press of the nominal force of 25 MN, in three forging operations (Fig. 2): operation I – upsetting, operation II – preliminary die forging and operation III – finishing forging (Fig 2b). The forged material is the C45 steel, in the shape of a cylinder. The data obtained from a detailed analysis of the lower die inserts from the first, second and third forging operation were entered into the system database (Fig 2d).

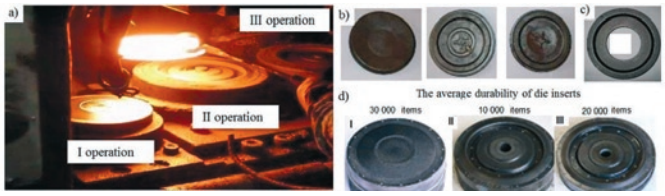


Fig. 2. The view of forging process: a) tools mounted on a press, b) forgings after consecutive operations, c) a ready front wheel forging (after punching and trimming, before toothing), d) lower die inserts for I, II, III operation

II Precision forging of a CVJB (constant velocity joint body) housing in closed dies – tools not preheated, lubricated and cooled, initial billet inductively heated to 920°C. The industrial forging of a CV universal joint housing in the GKN Driveline Forge consists of four hot forging operations and one cold forging operation (Fig. 3). The forging`s material is the XC45 steel. The dies and punches are made of the Unimax steel for hot working tools (1.2367). After the heat treatment tool have a hardness of 48-52 HRC. The tools are not nitrided and not heated before the process, are pre-charged

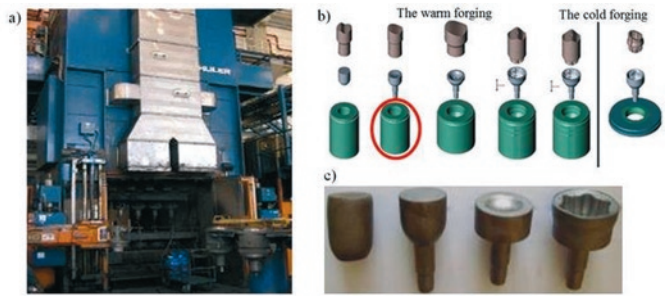


Fig. 3. The precision forging process: a) crank press, b) scheme of individual operations (V operation – marking), c) forgings after successive operations

to a special welding agent (Aerodag CERAMISHIELD) in order to reduce the risk of rupture in a thermal shock caused by contact with the hot preform at the beginning of process. In addition, the punch of the fourth operation is heated in a special device and 2-times immersed in a mixture of graphite with water (on the press is mounted, unheated). Tools after forging a few pieces forgings reach operating temperature (250-300° C).

III Forging of a fastener (type of Putanker 1,3T) used to move concrete slabs, in a closed die without lubrication and cooling tools, billet is heated inductively (locally, Fig 4b) to 1000 °C, in three operations on a PMS 160B eccentric press (Fig. 4a) with the use of a TR device (Fig. 4c). Figure 4d shows the forgings after each operation. Figure 4e shows the tools used in the first, second (Fig. 4f) and third (Fig. 4g) forging operation.

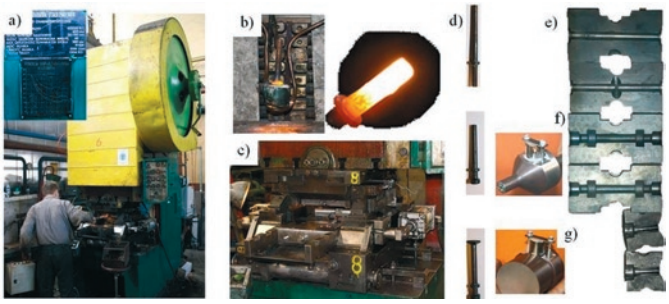


Fig. 4. The view of: a) PMS 160B eccentric press, b) heater with a locally heated preform for the second operation, c) TR device, d) forgings after successive operations, e) dies for the first operation, f) tools for the second operation (punch with AE sensor), g) tools for the third operation (punch with AE sensor)

The punches and the die inserts are made of steel WCL and ORVAR SUPREME (WCLV), which the hardness after heat treatment was 58 HRC. In addition, the second and third operations are used punches (in the first operation-upsetting is done by two dies, without the punch). The tools are not nitrided and pre-heated. Average life of the die inserts are: 12,000 for the first operation, for the second operation about 8000 forgings, and for the third operation about 10,000 forgings. The data obtained from a detailed analysis of only second die was entered into the system database.

After the selection of representative processes, their most important parameters were determined, such as: charge and tool temperature, forming force, velocities, tribological conditions, preform and final product geometry as well as tool shape. These parameters were determined based on the operation sheets. The remaining significant parameters, hard or impossible to determine (pressures, path of friction, temperatures in contact etc.) were determined by means of numerical modelling verified by a measuring and monitoring system [10]. The geometrical defect (wear) of the die impressions was determined based on a comparison of the superimposed laser scan images of the new and the worn tools (after a specific number of forgings) [6].

After the analysis of the selected die forging process, three groups of parameters were specified, which had a significant effect on the tool wear, that is:

- parameters of the forging process (number of forgings, number of operations, and for each operation, also: charge temperature; pressures; lubrication; cooling intensity etc.),
- parameters characterizing the tool (die) – Fig. 5a shows an exemplary schematic of the division of the die insert's cross section for the second operation of hot forging a front wheel in a closed die, while Fig. 5b presents the precision forging of a constant-velocity universal joint housing and Fig. 5c shows scheme of cross section of part die to forging fasteners. All forging tools are divided into elementary areas (1-9), described with symbols representing elementary shapes (A-F), distinguishable in any die.
- parameters characterizing the charge material.

The assumptions and system tasks (prediction of the life of the forging tools used in the hot die forging process, as well as identification of the critical points and wear mechanisms) made it possible to define the parameters providing the information on the tool wear, which are:

– failure mechanism (4 wear mechanisms were considered: thermo-mechanical fatigue, mechanical fatigue, abrasive wear and plastic strain) as well as the mechanism's participation in the failure of the given area;
– geometrical defect – loss of material (wear [mm]).

The detailed information on the process of obtaining primary data for the specified parameters describing the selected process was thoroughly discussed in [5]. Table 1 shows a fragment of the elaborated set of data.

The global table contains 450 knowledge vectors. The number of forgings was analyzed in the range of 0-25000.

4. The neural network model

The process of constructing a neural network model consists of the following stages: determination of independent and dependent variables; selection of neural network type and determination of its structure; neural network training and evaluation of network model.

4.1. Determination of independent and dependent variables

A neural network's task is to show the effect of the particular factors on the wear of the die during the die forging process. It calculates the participation of the particular mechanisms, i.e. thermo-mechanical fatigue, abrasive fatigue, plastic strain and mechanical wear, in the failure process. It also calculates the wear of the tool itself in the specified area.

Table 2. Dependent variables

L.P.	OUTPUT variables	Scope
1	thermo-mechanical fatigue	0-1
2	abrasive wear	0-1
3	plastic strain	0-1
4	mechanical fatigue	0-1
5	wear	-3 ÷ 0,05

Table 2 shows the dependent variables – the process parameters which are determined as the network output. As the independent variables, having an effect on the presented output variables, the following were selected from the global set of data: number of forgings, temperature, hardness, deformation time, total time, pressure,

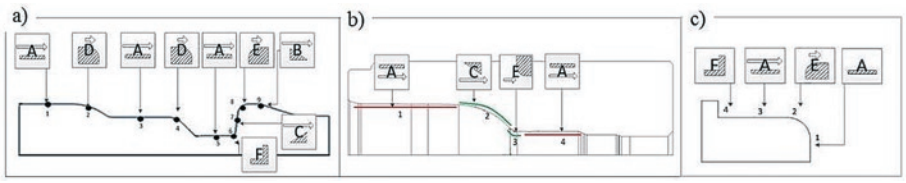


Fig. 5. Cross section with schematic division into areas and elementary shapes for: a) die insert used in forging front wheel, b) die used in precision forging of CVJB housing, c) a formed part of die for the first operation

Table 1. Fragment of table with data

number of forgings	shape	temperature [C]	nitrification [yes/no]	deformation time [s]	total time [s]	preasure [MPa]	path of friction	lubrication [1/0]	z-cm	z-s	o-p	z-m	wear [mm]
550	A	1200	no	0,162	1,663	160	średnia	0	0,50	0,50	0,00	0,00	-0,26
550	A	1100	no	0,162	1,663	160	średnia	0	0,00	1,00	0,00	0,00	-0,2
...	..	..	..	..	..	..	..	..	..	..	..	..	..
1850	E	1200	yes	0,038	0,558	500	very high	1	0,33	0,33	0,33	0,00	-0,4
1850	A	1200	yes	0,058	1,068	960	very high	1	0,60	0,20	0,20	0,00	-0,2
1900	A	1200	no	0,165	0,165	140	very high	0	0,30	0,30	0,40	0,00	-0,26
1900	A	1100	no	0,165	0,165	140	very high	0	0,30	0,40	0,30	0,00	-0,2
...	..	..	..	..	..	..	..	..	..	..	..	..	..
2200	E	1100	no	0,038	2,476	384	very high	0	0,00	0,50	0,00	0,50	-0,1
2200	E	1100	no	0,038	2,476	384	very high	0	0,00	0,50	0,00	0,50	-0,07
...	..	..	..	..	..	..	..	..	..	..	..	..	..
2500	A	1150	yes	0,088	0,088	1074	very high	1	0,35	0,35	0,30	0,00	-0,568
2500	E	1150	yes	0,087	0,087	1300	very high	1	0,33	0,33	0,33	0,00	-0,189
...	..	..	..	..	..	..	..	..	..	..	..	..	..
4000	E	1100	no	0,111	0,631	480	very high	1	0,25	0,25	0,30	0,20	-0,6
4000	C	1100	no	0,117	0,717	560	very high	1	0,33	0,33	0,33	0,00	-0,3
...	..	..	..	..	..	..	..	..	..	..	..	..	..
4300	E	1200	yes	0,038	0,558	500	very high	1	0,40	0,40	0,20	0,00	-1,6
4300	A	1200	yes	0,058	1,068	960	very high	1	0,40	0,30	0,30	0,00	-1,5
...	..	..	..	..	..	..	..	..	..	..	..	..	..
6000	A	1150	yes	0,088	0,088	1074	very high	1	0,35	0,35	0,30	0,00	-0,786
6000	A	1150	yes	0,072	0,072	620	very high	1	0,70	0,10	0,20	0,00	-0,065
...	..	..	..	..	..	..	..	..	..	..	..	..	..
6500	A	1100	no	0,021	1,134	320	very high	0	0,50	0,50	0,00	0,00	-0,19
...	..	..	..	..	..	..	..	..	..	..	..	..	..
6900	E	1200	yes	0,038	0,558	500	very high	1	0,40	0,40	0,20	0,00	-2,1
...	..	..	..	..	..	..	..	..	..	..	..	..	..
7044	C	920	no	0,117	0,717	560	very high	1	0,80	0,20	0,00	0,00	-0,01
7044	A	920	no	0,097	0,717	650	very high	1	0,85	0,10	0,00	0,05	0
...	..	..	..	..	..	..	..	..	..	..	..	..	..
8000	E	1100	no	0,111	0,631	480	very high	1	0,25	0,25	0,30	0,20	-1,5
...	..	..	..	..	..	..	..	..	..	..	..	..	..
9000	A	1200	no	0,162	1,663	160	very high	0	0,30	0,30	0,40	0,00	-2,6
9500	E	1200	yes	0,038	0,558	500	very high	1	0,33	0,33	0,33	0,00	-2,7
...	..	..	..	..	..	..	..	..	..	..	..	..	..
12000	C	1100	no	0,117	0,717	560	very high	1	0,33	0,33	0,33	0,00	-2,4
12000	E	1100	no	0,111	0,631	480	very high	1	0,25	0,25	0,30	0,20	-2,4
...	..	..	..	..	..	..	..	..	..	..	..	..	..
12750	D	1150	yes	0,087	0,087	941	very high	1	0,45	0,45	0,10	0,00	-1,285
12750	A	1150	yes	0,072	0,072	620	very high	1	0,80	0,10	0,10	0,00	-0,161
12750	D	1150	yes	0,063	0,063	542	very high	1	0,90	0,05	0,05	0,00	-0,032
...	..	..	..	..	..	..	..	..	..	..	..	..	..
15500	A	1200	no	0,162	1,663	160	very high	0	0,30	0,30	0,40	0,00	-2,6
15500	A	1100	no	0,162	1,663	160	very high	0	0,30	0,40	0,30	0,00	-2,1
16000	A	1100	yes	0,084	0,604	210	very high	1	0,33	0,33	0,33	0,00	-0,8
17054	A	920	no	0,097	0,717	650	very high	1	0,60	0,25	0,00	0,15	-0,05
17054	C	920	no	0,117	0,717	560	very high	1	0,60	0,25	0,00	0,15	-0,04
22320	A	920	no	0,097	0,717	650	very high	1	0,70	0,15	0,00	0,15	-0,04
22320	A	920	no	0,084	0,604	210	very high	1	0,65	0,15	0,00	0,20	-0,01
25000	B	1100	yes	0,005	1,444	30	very high	1	1,00	0,00	0,00	0,00	-0,7
25000	E	1200	yes	0,034	1,444	800	very high	1	0,40	0,30	0,30	0,00	-0,22
25000	E	1100	yes	0,034	1,444	800	very high	1	0,80	0,20	0,00	0,00	-0,17
25000	D	1200	yes	0,018	1,444	600	very high	1	0,70	0,20	0,10	0,00	-0,13

path of friction, lubrication and shape. Table 3 shows the selected independent variables (neural network inputs).

4.2. Neural network structure and parameters

The set of data describing the modelled phenomenon (450 knowledge vectors), was divided into the training set (70%), the validation set (15%) and the test set (15%). At the network training stage, the training set is applied. The validation set is used to control the course of training by way of verifying how well the neurons are trained. In practice, the training included two phases: selection of weights for the training set and weight testing on samples from the validation set. The modification of the weight values continued until approximation error minimization was achieved or the error in the validation set began to grow.

For the determination of the detailed, as well as optimal, network architecture, the STATISTICA program and its module *Automatic Neural Networks* were used. Tests were performed on a few tens of architectures with different numbers of hidden neurons and different activation functions in the hidden and the output layer (linear, sigmoidal (logistic), tangensoidal and exponential). From among all the networks generated by the program, the one characterizing in the low-

Table 3. Idependent variables

L.P.	INPUT variables	Scope
1	number of forgings	0-25000
2	temperature [°C]	900-1200
3	nitrification	yes/no
4	deformation time [s]	0-0,185
5	total time [s]	0-2,5
6	pressure [MPa]	0-1300
7	path of friction	low, medium, high, very high
8	lubrication	yes/no
9	shape	A, B, C, D, E, F

est validation error (at the level of 12%) was ultimately selected. An additional measure of the model quality was the Pearson linear correlation coefficient (*R*), calculated in the particular set types (training, validation and test), for the network and the set data reply. The correlation coefficient for the training set reached the value of 0,937, for the validation set - 0,828 and for the test set - 0,842. The assumed error function is the sum of squared deviations (SOS) between the set value and network output calculated for each set. The error value for the training set is at the level of 3%, for the validation set - 12%, and for the test set - 9%. Considering the problem's degree of complication, the number of input data and the assumed outputs, it can be stated that the network parameters are at a sufficient level. Hyperbolic tangent (Tanah) was assumed as the activation function in the hidden

Table 4. Parameters of MLP 19-25-5 neural network

Name of network	MLP 19-25-5
Error (training)	0,039309
Error (validation)	0,127667
Error (testing)	0,090174
Quality (training)	0,937455
Quality (validation)	0,828000
Quality (testing)	0,842213
Training algorithm	BFGS 56
Error function	SOS
Activation (hidden)	Tanh
Activation (output)	Logistic

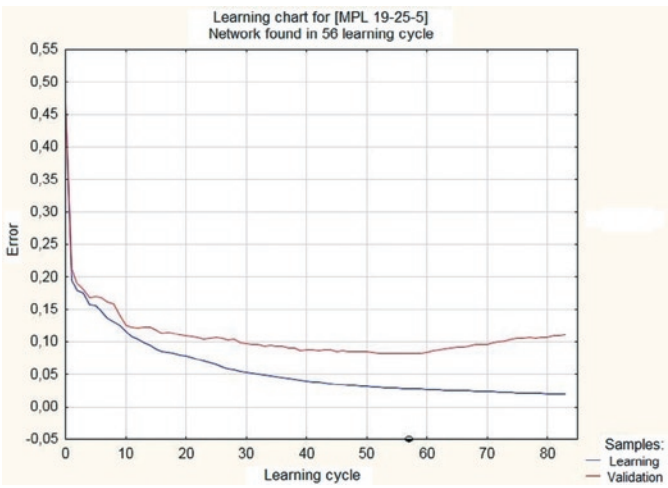


Fig. 6. Changes in errors (training and validation) depending on epoch number for MLP 19-25-5 network

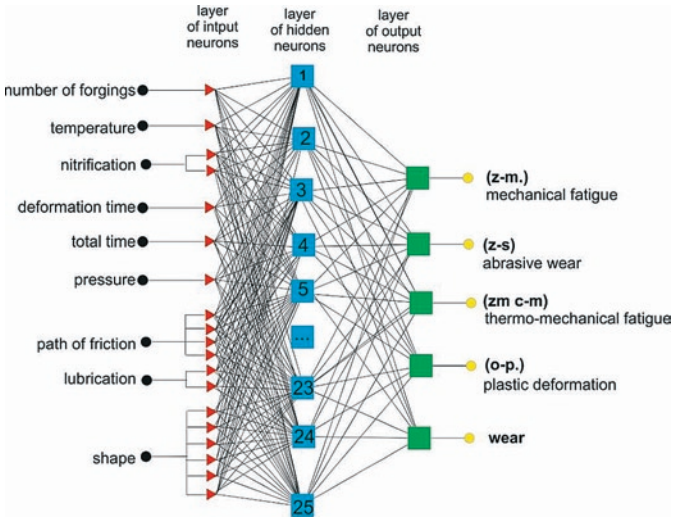


Fig. 7. Diagram of the elaborated MLP 19-25-5 network

neuron layer, and the logistic (sigmoidal) activation function was assumed for the layer of output neurons. Table 4 shows the basic parameters of the elaborated network.

In the case of the selected MLP 19-25-5 network, the assumed minimal approximation error was not achieved; the training process was terminated in 56 epoch, when the validation error started to grow. In this network, the BFGS (Broyden-Fletcher-Goldfarb-Shanno) method was used for training. Fig.6 shows the process of change in the error rate (training and validation) during training.

A simplified structure diagram of the selected network is presented in Fig. 7. The network is of the MLP type, consisting of one input layer (9 input variables/19 neurons), one hidden layer (25 neurons) and one output layer (5 neurons).

4.3. Error analysis

Table shows the basic quality parameters for the elaborated network. The given errors (training, validation, test) concern the total network error calculated as the mean value of the remainder squares for all five output variables.

A detailed analysis of the errors for the MLP19-25-5 network, for the particular output variables, i.e. wear (**wear**) thermo-mechanical fatigue (**z-cm**), abrasive wear (**z-s**), plastic strain (**o-p**), mechanical fatigue (**z-m**), separated according to the particular sets: training, validation, test, are compiled in Tables 5-9. The tables contain the following characteristics: remainder squares mean value, mean absolute error, correlation coefficient.

In the analysis of Tables 5-9, we can notice that the elaborated network provides the best results for the mechanical fatigue output variable (z-m). The correlation coefficient for the test set is at the level of 0,99, which gives a very high linear approximation, and the mean absolute error equals 0,01. These results are presented in Table 6. The mean absolute error at the level of 0,04, for the test set, is exhibited by the plastic strain output variable (o-p), presented in Table 7, where the correlation coefficient for the test set equals 0,8. The lowest approximation and the highest error can be expected for the thermo-

Table 5. Summary of the network fit level for the observed variable wear

wear	Remainder squares mean value	Mean absolute error	Correlation coefficient
Training	0,0306	0,1175	0,9577
Validation	0,1161	0,1833	0,8147
Testing	0,0573	0,1579	0,9200

Table 6. Summary of the network fit level for the observed variable z-m

z-m	Remainder squares mean value	Mean absolute error	Correlation coefficient
Training	0,0044	0,0196	0,9664
Validation	0,0039	0,0179	0,9759
Testing	0,0009	0,0114	0,9954

Table 7. Summary of the network fit level for the observed variable o-p

o-p	Remainder squares mean value	Mean absolute error	Correlation coefficient
Training	0,0018	0,0221	0,9425
Validation	0,0043	0,0366	0,8992
Testing	0,0061	0,0490	0,8077

Table 8. Summary of the network fit level for the observed variable z-s

z-s	Remainder squares mean value	Mean absolute error	Correlation coefficient
Training	0,0114	0,0659	0,9263
Validation	0,0359	0,1051	0,7711
Testing	0,0356	0,1122	0,7656

mechanical fatigue output variable (z-cm), for which the correlation coefficient for the test set is at the level of 0,72, and the mean absolute error equals 0,17 (table 9).

Table 9. Summary of the network fit level for the observed variable z-cm

z-cm	Remainder squares mean value	Mean absolute error	Correlation coefficient
Training	0,0303	0,1064	0,8943
Validation	0,0949	0,1772	0,6790
Testing	0,0802	0,1722	0,7222

Conducted by the authors of the study in the scope of tool life (material and others [11, 12]), concerning the input variables and their sequence, are correct and the obtained hierarchy of importance is mostly determined by the amount of data and its type. Only the case of the pressure variable, for the elaborated network, is somewhat thought-provoking, as in the case of abrasive wear, described by means of the Archard model, the volume of wear is proportional to i.a. pressure. On the other hand, the Archard model proves effective mainly for non-lubricable contacts. What is more, as it was pointed out earlier, the input variables are generally not independent, and so, even experienced researchers are faced with doubts when separately analyzing some of the results.

4.4. Modeling results analysis

The analysis of the results obtained in the inference process, with the use of a virtual experiment, as compared with the results obtained empirically, shows an error at the level of 0-10%. Considering the fact that, in the actual industrial processes, such differences can assume even higher values for the particular tools, and the determination of the dominant mechanism is not always possible, even when it is performed by an experienced operator, the proposed inference model should be treated as sufficient for the analysis of the described process. Exemplary inference results obtained with the use of the elabo-

Table 10. Inference results for selected input scenario

1	2	3	4	5	6	7
INPUT	**number of forgings**	550	1850	1850	4000	4300
	temperature	950	950	1200	1100	950
	deformation time	0,183	0,048	0,048	0,083	0,027
	total time	1,684	1,068	1,068	0,603	1,444
	pressure	168	640	640	210	400
	lubrication	0	1	1	1	1
	shape	A	D	D	A	E
	nitrification	no	yes	yes	no	yes
	path of friction	low	low	low	high	medium
OUTPUT MLP19-25-5	**wear**	0,0493	0,0497	0,0342	-0,0182	-0,105
	z-cm	0,000	0,877	0,912	0,403	0,149
	z-s	0,000	0,113	0,028	0,295	0,836
	o-p	0,000	0,000	0,059	0,300	0,013
	z-m	0,000	0,010	0,001	0,000	0,000
OUTPUT EXPERIMENTAL DATA	**wear**	0,000	0,000	-0,040	-0, 020	-0,250
	z-cm	0,000	1,000	1,000	0,333	0,000
	z-s	0,000	0,000	0,000	0,333	1,000
	o-p	0,000	0,000	0,000	0,333	0,000
	z-m	0,000	0,000	0,000	0,000	0,000

rated neural network for selected input data scenarios, are presented in Table 10.

The data presented in Table 10 confirms a good agreement and a correct tendency of the calculations obtained from the elaborated network, in reference to the experimental test results for the set values of input parameters. The differences are at the level of 10%, which is confirmed by the error set to evaluate the network. For example, for the input values presented in column 6 (number of forgings = 4000, temperature = 1100, deformation time = 0,083, total time = 0,63, pressure = 210, lubrication = 1, shape = A, nitrification = no, path of friction = high), the network demonstrated wear at the level of -0,0182 mm, that is very close to 0, which is confirmed by the experimental tests. According to the network, the wear mechanisms for this input scenario are: thermo-mechanical fatigue (z-cm) at the level of 0,403, abrasive wear (z-s) = 0,295 and plastic strain (o-p) = 0,300. The participation of the mechanical fatigue mechanism (z-m) was determined by the network to be at the level of 0. The experimental data in this case confirms the correctness of the results provided by the network. Precisely, these three mechanisms, with the participation of 0,33 each, were pointed to in the experimental tests. Fig. 8 illustrates an interface system developed which allowed for testing of the model.

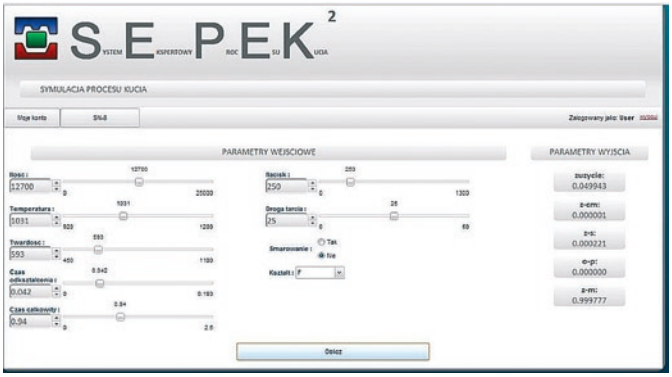


Fig. 8. Presentation of calculations by means of SEPEK interface

The developed interface is very user-friendly, as it allows for an intuitive and efficient determination of the values of the input variables (by way of shifting the slide or directly entering the data). The determination of any interesting values of input variables (number of forgings, charge temperature, hardness, deformation time and total time, pressures, path of friction, lubrication and the representative elementary tool shape), automatically generates a report on the output parameters. Preliminary tests of its use by employees working in selected die forges (i. a.: Kuźnia Jawor, Kuźnia Polska) showed a high ease into use and high functionality and practical usability.

4.5. Global interesting results

After successfully verifying the actions a decision support system authors conducted a a global analysis using ANN. Fig. 9a shows an example of the results of the percentage share the main mechanisms of destruction and loss of geometric material for "common conditions" prevailing in the industrial processes of forging, in the case of tools with lubrication. In contrast, in Fig. 9b results for forging processes, that do not use cooling and lubricant. Table 11 shows the most important parameters for the two variants of working tools: a) with lubrication and cooling, and b) without lubrication and cooling.

The results indicate that in the hot forging process for the heated tool (nitriding) and cooled, the dominant, destructive mechanism (for average number of forgings of about 15,000) is a thermo-mechanical fatigue, rather than abrasive wear. These results indicate a slightly different interpretation accepted in literature for the common view of the fact, that statistically 70% of the forging dies is withdrawn from production due to the loss dimensions - due to abrasive wear and plastic deformation, 25% – as a consequence of fatigue cracks and only 5% for other reasons (non-compliance with the technology, construction and material defects, thermal and thermo-chemical treatment defects, etc.). For a case when the tools are not lubricated and cooled percentage of the abrasive wear is dominant, increases also the share of plastic deformation. Conducted research and analyzes for other values characterizing parameters of the tools (pressure, contact times and the deformation temperature of the initial billet, etc.) confirm the relationship shown the dominance of thermal-mechanical fatigue in relation to "easily measurable" abrasive wear. In contrast, it was also observed that in case the tools are lubricated, but do not have the protective layer in the form of "nitriding" thermo-mechanical fatigue is also a parent, wherein, with the number of forging (the average number of about 7000-8000 forgings), much faster, the predominant destructive mechanism is abrasive wear.

The results presented in the diagrams were obtained based on the experimental data (450 records of knowledge were elaborated), which were used to construct a decision support system based on ASN for the analysis and prediction of forging tool durability. On this basis, the system generated the results for two different sets tool operation conditions (for exemplary work conditions data, shown in Table 11).

In the case of diagram b), Fig. 9, the collection of experimental data (vectors of knowledge) also included the values of material growth for a small number of forgings, which, as confirmed by the studies, was the result of adhesion of the forging and tool material, hence the positive values.

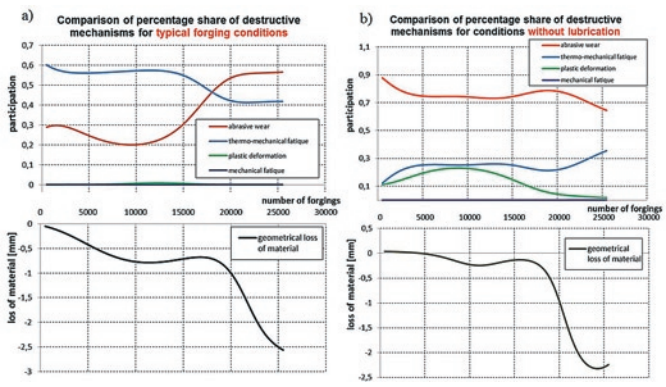


Fig. 9. Comparison of the percentage share *results* of the *main mechanisms of destruction and loss of material* with *growing number of forgings: a) typical conditions - tools heated, lubricated and cooled, b) "non-standard conditions" prevailing in the industrial processes of forging - without lubrication*

Table 11. The most important parameters of the forging tools: a) the typical conditions prevailing in the forging process for *tools* with *lubrication, b) tools without lubrication and cooling*

Work conditions	Temperature of billet [°C]	Normal stress of Surface tool detemined by FEM [MPa]	Forming time (forging) [s]	Total contact time deformated material with tool [s]	Nitriding	Lubrication and cooling
With lubrication	1100	650	0,097	0,717	yes	yes
Without lubrication	1100	650	0,097	0,717	no	no

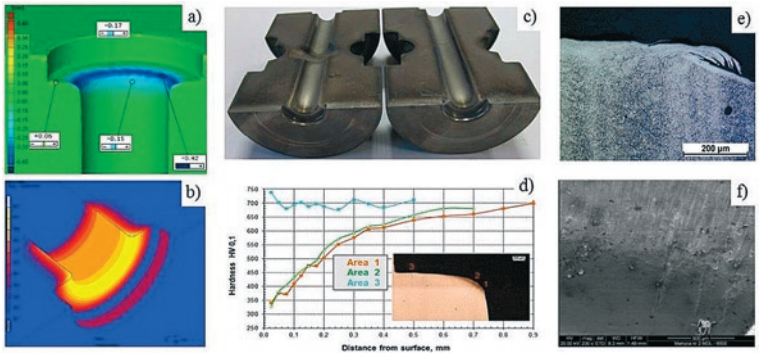


Fig. 10. Analysis of non-lubricated and non-cooled tools: a) results of scanning for the collar area, b) temperature distribution in the most worn area, obtained from MES, c) view of the dies used to forge a construction catch in the first operation, d) microhardness measurement results, e) plastic deformations in the analyzed area – metallographic microscope, e) traces of abrasive wear – scanning microscope

The local extremes in both diagrams with the number of over 15 thousand forgings, are caused by the fact that, in the experimental studies, different tools from the same processes were selected (after different numbers of produced forgings), and so, the obtained results e.g. of the measurement of material loss for the tool after 12 thousand forgings and another after 15 thousand, can vary slightly.

It should be noted that these results, obtained with the use of the elaborated system, provide error at the level of up to 10%. Also, the experimental data did not include (cover) all the areas, i.e. from 0 to 25 thousand forgings.

What is more, it was observed that, due to the working conditions of the tools with similar numbers of forgings (about 17000, for both sets of tool working conditions), a rapid increase of material loss occurred. The macro- and micro-studies showed that larger parts of material were detached, e.g. of the nitrided layer from different areas or the oxidized and cracking network of thermo-mechanical fatigue, which intensified the mechanism of abrasive wear, as they worked as hard particles, abrading and taking away other parts of the tool material. It is interesting to notice that, in the case of lubricated and cooled tools, from as few as 20 thousand forgings up, we observe an increase of the participation of thermal fatigue in respect of abrasive wear. This case is slightly different than the one of non-cooled and non-lubricated tools, for which, with the same number of forgings, a drop of abrasive wear was observed, with a simultaneous increase of thermo-mechanical fatigue.

The studies of the methods of evaluation and analysis of forging tool durability, presented in works [8, 9, 11], have shown that: most of the degradation mechanisms in the die forging processes performed at elevated temperatures cause and reveal themselves in the form material loss or shape change. The wear intensity changes with the change of the process parameters, which is determined mainly by the contact time, pressure values, temperature changes and tribological conditions. That is why it has been commonly assumed that it is abrasive wear which is the dominating degradation mechanism, whereas the second dominating mechanism is plastic deformation. And so, very often, in many elaborations, most of the degradation mechanisms are modelled by means of the Archard abrasive wear model. As it has been demonstrated by the results [6, 8] of the presented microstructural tests, the results of surface scanning (Fig 10-11) and implemented to database in the decision support system and the results obtained from the expert system, the most frequently occurring as well as dominating mechanism is thermo-mechanical fatigue, which additionally accelerates the other degradation mechanisms. On this basis, it can draw, that in hot die forging processes, in the case when the tools, in a particular operation or in the whole process, are not lubricated and cooled, it can be assumed that abrasive wear is the actual dominant degradation mechanism, and thermo-mechanical fatigue plays a secondary role. In such cases, the percentage of plastic deformations increases as well, as a result of the effect of temperature, which causes local tempering of the tool material and lowering of durability. Fig. 10 shows images of exemplary tools, in the analyzed processes, for which the descried situation takes place.

In contrast, for forging tools working under typical conditions, such as hot and semi-hot die forging processes, which are lubricated and cooled, thermo-mechanical fatigue is mostly the dominant mechanism. As a result of periodical temperature changes, during the forging process, we observe an alternating expansion and 'shrinkage' of the surface layer of the die, which, in consequence, leads to the formation of a thermal crack network. The latter, as a result of periodical mechanical loads, causes an increase in the stress concentration and expands by forming a primary and, in time also, secondary crack network. The spalling of the thermo-mechanical fatigue network causes a further expansion of cracks and intensifies the abrasive wear. The crack expansion is also favoured by the presence of scale, which, by filling the crack, can work as a 'wedge'. Under such process conidtions, plastic deformations basically do not occur. Fig. 11 presents a comparison of the state of the tools (die inserts) after performance, used in the industrial process of forging a front wheel, for the first operation (no lubrication or cooling) as well as for the second operation (the tools are lubricated and cooled).

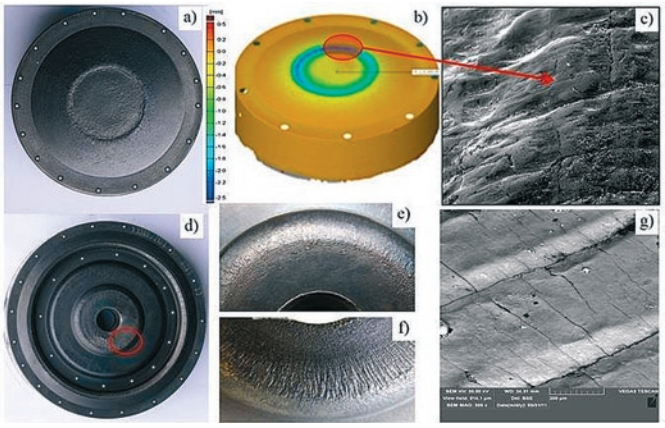


Fig. 11. Comparative analysis of unlubricated tools with lubricated and cooled tools, used in the same forging process: a) lower die insert after 9000 forgings, b) tool scanning results, c) SEM image from the selected die area – traces of plastic deformation, abrasive wear and oxidation, d) lower die insert after 1850 forgings, e) image of the insert front from the selected area after 550 forgings, f) after 4300 forgings, g) SEM image of the insert front from the marked area in Fig. 11d – visible primary network of thermal fatigue, traces of grooves formed as a result of abrasive wear by hard particles

Fig. 11 shows the results of the state of tools (die inserts) after their exploitation, used in the manufacturing process of wheel forging process for different working conditions confirmed the complexity of the occurrence of many destructive mechanisms, both in different regions of the tool, as well as the change and the progression of wear with increasing the number of forgings.

The presented comparison illustrates very well the complexity and variety of the mechanisms occurring in the case of forging tools, which confirms my conviction even more that the analysis of durability is still an existing challenge.

5. Summary

The work involves the use of artificial neural networks in a decision support system in durability prediction of forging tools used in the hot die industrial forging processes.

Forging tools (dies and punches) designed by engineers and technologists used in the die industrial forging process wear usually much earlier than it was planned. The durability of forging tools depends primarily on: the conditions under which extends the forging process, the design and construction of tools, their proper heat treatment, appropriate for the tool material, the shape initial billet and preform and also a lubrication and cooling systems, etc. Therefore, it is advisable to systems development systems which provide the best choice of tools working conditions in order to increase its life. Developed a decision support system is an excellent IT, which could support and supplement work of process engineer in the selection of the optimal working conditions of forging tools. In the manuscript presented a prototype version of such a system.

Presented in the article the system has been developed based on the experimental data obtained for selected representative industrial die forging processes, which include most of the processes implemented in die forges. As the formal tool representing the knowledge in the system, an artificial neural network was used. The set of data applied for network training contained 450 training cases coming from the performed experimental research as well as computer simulations. The results show that, by parameterizing the crucial factors of the forging process, it is possible to develop an evaluation system of the percentage of the typical failure mechanisms (thermo-mechanical fatigue, mechanical wear, abrasive wear and plastic strain) and to calculate the value of the geometrical defect of the tools.

Predicting the degree of wear/failure of the die and the type of mechanisms responsible for this, with the assumed parameters of its work, is a very complex process, difficult to design. The collected source data and the neural network elaborated on its basis make the assessment error at the level of 10%, which, considering the complication of the problem, is a satisfactory result. The level of the global error with which the network model performs the calculations is at the level of 10% (testing error = 0,09) and it is comparable with the value of the error determined for the first version of the SEPEK system, formalized by means of fuzzy logic (0-10% for the percentages of wear mechanisms; 0-15% for the parameter of the wear degree). Considering the large differences in the wear of the particular tools and the fact that pointing to the mechanism determining the wear cannot always be clear, even in the case when it is performed by an experienced expert, it can be assumed that the elaborated system provides results with an acceptable error. One should also emphasize the fact that the results obtained by the system have been verified and positively evaluated by experts.

Carried out thanks to decision support system-based on ANN global analysis and other durability testing showed that fatigue thermo-mechanical destruction due to oxidation, very often occur together with the mechanism of wear, creating a synergy effect, causing the acceleration, the most visible and "easily measurable" process abrasive wear. So, as clearly mentioned in the available literature rightly given that as many as 70% of all withdrawn of forging tools from further exploitation is the result of wear. But not informed by the fact that a significant part of this share is caused by the strengthening of the destructive mechanism mainly due to thermo-mechanical fatigue. The synergy between these mechanisms results in detachment of non-cyclic large particles for tool with a primary or secondary grid of cracks and detachment of the cyclic much smaller particles - scale in the form of hard oxides. All of these particles as a result of thermal fatigue and oxidation work as a kind of abrasive greatly intensifying the destruction process as a result of wear (abrasive wear). This in turn leads to a sometimes very large changes in geometry tools which translate directly to the forging, which from the point of view quality and functionality of such a product is not permitted.

The further work aiming at perfecting the model will be connected with the process of optimizing the network and introducing a larger amount of training data obtained from the consecutive experimental data.

The results obtained and presented in the manuscript are distinctly application character, because based on the these analysis concerning of destructive mechanisms can be used appropriate methods or preventive measures that will allow to increase the durability of forging tools. The recipients of this type of decision support systems are, as revealed preliminary research especially technologists and engineers working in the die forges.

Acknowledgements

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

References

1 Altan T. Cold and hot forging fundamentals and application. ASM Internation, Ohio 2005.
2. Azari A, Poursina, M. i Poursina, D. Radial forging force prediction through MR, ANN and ANFIS models. Neural Computing & Applications 2014; 25(3-4): 849-858, https://doi.org/10.1007/s00521-014-1562-8.
3. Gangopadhyay T, Kumar D, Pratihar I. Expert system to predict forging load and axial stress. Appl. Soft. Comput. 2014; 11(1): 744-753, https://doi.org/10.1016/j.asoc.2009.12.036.
4. Gronostajski Z, Hawryluk M, Jaśkiewicz K, Niechajowicz A, Polak S, Walczak S, Woźniak A. Application of physical and mathematical modelling to analysis of different forging processes of constant velocity joint body. Computer Methods in Materials Sciences 2007; 7(2): 231-236.
5. Gronostajski Z, Hawryluk M, Kaszuba M, Marciniak M, Niechajowicz A, Polak S, et al. The expert system supporting the assessment of the durability of forging tools. The International Journal of Advanced Manufacturing Technology 2015; 82(9): 1973-1991.
6. Gronostajski Z, Hawryluk M, Kaszuba M, Ziemba J. Application of a measuring arm with an integrated laser scanner in the analysis of the shape changes of forging instrumentation during production. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (2): 194-200, https://doi.org/10.17531/ein.2016.2.6
7. Gronostajski Z, Hawryluk M. The main aspects of precision forging. Archives of Civil and Mechanical Engineering 2008; 8(2): 39-57, https://doi.org/10.1016/S1644-9665(12)60192-7.
8. Gronostajski Z, Kaszuba M, Hawryluk M, Zwierzchowski M. A review of the degradation mechanisms of the hot forging tools. Archives of Civil and Mechanical Engineering 20014; 14: 528-539, https://doi.org/10.1016/j.acme.2014.07.002.
9. Gronostajski Z, Kaszuba M, Polak S, Zwierzchowski M, Niechajowicz A, Hawryluk M. The failure mechanisms of hot forging dies. Materials Science and Engineering. A, Structural Materials: Properties, Microstructure and Processing 2016; 657: 147-160, https://doi.org/10.1016/j.msea.2016.01.030.
10. Hawryluk M, Kaszuba M, Gronostajski Z, Sadowski P. Systems of supervision and analysis of industrial forging processes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (3): 315-324, https://doi.org/10.17531/ein..2016.3.1

11. Hawryluk M. Metody analizy oraz zwiększania trwałości narzędzi kuźniczych stosowanych w procesach kucia matrycowego na gorąco. Monograficzna seria wydawnicza Problemy Eksploatacji i Budowy Maszyn, ISBN 978-83-7789-410-1, Wyd. Naukowe ITE - PIB, Radom 2016.
12. Hawryluk, M. Review of selected methods of increasing the life of forging tools in hot die forging processes. Archives of Civil and Mechanical Engineering 2016;16: 845-866, http://dx.DOI:0.1016/j.acme.2016.06.001.
13. Heinemeyer D. Gensekschäden und Einflussgrössen der Standmenge. Industrieanzeiger 1978; 100 (73).
14. Katayama T, Akamatsu M, Tanaka Y. Construction of PC based expert system for cold forging process design. J Mater. Process. Technol. 2004; 155-156: 1583-1589, https://doi.org/10.1016/j.jmatprotec.2004.04.256.
15. Kocańda S, Kocańda A. Niskocyklowa wytrzymałość zmęczeniowa metali. PWN, Warszawa 1989.
16. Lange K, Cser L, Geiger M, Kals J,A.G. Tool Life and Tool Quality in Bulk Metal Forming. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture November 1993; 207 : 223-239, https://doi.org/10.1243/PIME_PROC_1993_207_085_02.
17. Lapovok R, Smirnov S, Shveykin V. Damage mechanics for the fracture prediction of metal forming tools. International Journal of Fracture 2000; 103(2): 111-126, https://doi.org/10.1023/A:1007593623392.
18. Li M, Liu,X, Xiong A. Prediction of the mechanical properties of forged TC11 titanium alloy by ANN. Journal Of Materials Processing Technology 2012; 121(1): 1-4, https://doi.org/10.1016/S0924-0136(01)01006-8.
19. Mazurkiewicz D. Maintenance of belt conveyors using an expert system based on fuzzy logic. Archives of Civil and Mechanical Engineering 2015;15(2) : 412-418, https://doi.org/10.1016/j.acme.2014.12.009.
20. Osowski, S. Sieci neuronowe w ujęciu algorytmicznym. WNT. Warszawa 1996.
21. Rauch L, Chmura A, Gronostajski Z, Pietrzyk M, Zwierzchowski M. Cellular automata model for prediction of crack initiation and propagation in hot forging tools. Archives of Civil and Mechanical Engineering 2016; 16(3) : 437-447, https://doi.org/10.1016/j.acme.2016.02.008.
22. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review 1958; 65(6): 386- 408, https://doi.org/10.1037/h0042519.
23. Smolik J. Rola warstw hybrydowych typu warstwa azotowana/ powłoka PVD w procesie zwiększania trwałości matryc kuźniczych. WITE. Radom 2007.
24. Subba Rao A.V, Pratihar D,K. Fuzzy logic-based expert system to predict the results of finite element analysis. Knowl-Based Syst. 2007; 20: 37-50, https://doi.org/10.1016/j.knosys.2006.07.004.
25. Sun Y, Hu L. Modelling optimisation of hot processing parameters of Ti-6Al-4V alloy using artificial neural network and genetic algorithm. Materials Research Innovations 2014; 18: 1052-1056, https://doi.org/10.1179/1432891714Z.000000000856.
26. Tadeusiewicz R, Gąciarz T, Borowik B, Leper B. Odkrywanie właściwości sieci neuronowych, przy użyciu programów w języku C #. Polska Akademia Umiejętności, Międzywydziałowa Komisja Nauk Technicznych, Kraków 2007.
27. Tadeusiewicz R. Neural networks in mining sciences - general overview and some representative examples. Arch. Min. Scien. 2015; 60 (4): 971-984, https://doi.org/10.1515/amsc-2015-0064.
28. Tadeusiewicz R. Neural networks.PWN. Warszawa 1993.
29. Tompos A, Margitfalvi JL, Tfirst E, He´berger K. Predictive performance of 'highly complex' artificial neural networks. Appl Catal Gen. 2007; 324: 90-93, https://doi.org/10.1016/j.apcata.2007.02.052.

Marek HAWRYLUK
Department of Metal Forming and Metrology
Faculty of Mechanical Engineering
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego 25, 50-370 Wroclaw, Poland

Barbara MRZYGŁÓD
Department of Applied Computer Science and Modelling
Faculty of Metals Engineering and Industrial Computer Science
AGH – University of Science and Technology
Al. Mickiewicza 30, 30-059 Cracow, Poland

E-mails: marek.hawryluk@pwr.edu.pl, mrzyglod@agh.edu.pl

Article citation info:
WANG Y, BI L, WANG S, LIN S, XIANG W. The application of dynamic bayesian network to reliability assessment of emu traction system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 349–357, http://dx.doi.org/10.17531/ein.2017.3.5.

Yanhui WANG
Lifeng BI
Shujun WANG
Shuai LIN
Wanxiao XIANG

THE APPLICATION OF DYNAMIC BAYESIAN NETWORK TO RELIABILITY ASSESSMENT OF EMU TRACTION SYSTEM

ZASTOSOWANIE DYNAMICZNYCH SIECI BAYESOWSKICH DO OCENY NIEZAWODNOŚCI ELEKTRYCZNEGO SYSTEMU TRAKCYJNEGO

The article introduces a novel application of a Dynamic Bayesian Network (DBN) in the reliability assessment with regard to the traction system of Electric Multiple Units (EMU), which focus on modeling approach to DBN construction. As a result of high complexity and growing interdependencies, it is increasingly vulnerable to the failure of components. Although many studies on the use of BN for estimating the system reliability have been conducted, there is a lack of effective modeling power regarding current tools in depicting both functional and temporal dependencies between components. In this paper, a new modeling approach to DBN generation is submitted, which can be applied to the system made up of certain components and different types of flows propagating through them. The Component-based CPT (Conditional Probability Table) and Time-dependent CPT are used to describe functional dependencies and temporal dependencies respectively. As the complexity of the system cannot be modeled in a tractable way as a DBN, a Breadth-First-Search (BFS) algorithm is introduced for the construction of the DBN model in an automated manner. With the application of the proposed DBN-based approach, the reliability of the traction system can be evaluated at any given time, which is of great significance to determine the plan of maintenance in an effort to ensure the system safety.

Keywords: *DBN; traction system; reliability assessment; breadth-first-search algorithm.*

W artykule omówiono nowatorskie zastosowanie dynamicznej sieci bayesowskiej (DBN) do oceny niezawodności elektrycznego systemu trakcyjnego ze szczególnym uwzględnieniem metod modelowania DBN. W związku z rosnącą złożonością elektrycznych systemów trakcyjnych oraz wynikającą z niej coraz większą ilością współzależności między komponentami, systemy te narażone są coraz częściej na awarie części składowych. Chociaż istnieje wiele badań dotyczących oceny niezawodności systemów trakcyjnych, stosowane obecnie narzędzia nie mają odpowiedniej mocy modelowania koniecznej do opisu zależności funkcjonalnych i czasowych pomiędzy częściami składowymi. W niniejszej pracy zaproponowano nową metodę modelowania generowania DBN, którą można stosować w odniesieniu do systemów składających się z pewnych określonych komponentów oraz różnych typów rozchodzących się przez nie przepływów. Zależności funkcjonalne i czasowe opisano, odpowiednio, za pomocą tablicy komponentowych prawdopodobieństw warunkowych (Component-based Conditional Probability Table, CPT) oraz tablicy czasowo-zależnych prawdopodobieństw warunkowych. Ponieważ złożoność systemu nie pozwala na zamodelowanie go w prosty sposób jako DBN, do automatycznej budowy modelu DBN wykorzystano algorytm przeszukiwania wszerz (Breadth-First-Search). Oceny niezawodności systemu trakcyjnego z wykorzystaniem proponowanej metody opartej na DBN można dokonywać w dowolnym czasie, co ma ogromne znaczenie przy planowaniu konserwacji w celu zapewnienia bezpieczeństwa systemu.

Słowa kluczowe: *dynamiczna sieć bayesowska, system trakcyjny, ocena niezawodności, algorytm przeszukiwania wszerz.*

1. Introduction

In recent years, a new era has seen the development of high-speed railway in China. By the end of 2012, China has boasted the coverage of about 9,356-km-high-speed railway[11]. The Beijing-Shanghai High-speed Railway which began to operate in July, 2011 has further pushed China towards super-high-speed trains with an operating speed of 380 km/h [20]. During the 13th five-year plan, the high-speed railway is supposed to increase to 30,000km, covering more than 80% of big cities. This widespread coverage has definitely rendered the reliability of EMUs a top priority. Nowadays, EMUs are generally ascribed to the extreme complexity and interdependencies as a result of the systematic use of new technologies (such as artificial intelligence, information/communication technologies, or communication networks). Failures of EMUs could cause a catastrophic accident, for example, the Wenzhou High-speed train crash on July 23, 2011. To sum up, the extreme reliability, the most critical of EMUs regarding the traction system, can never be underestimated.

Over the past decade, the need to conduct an analysis of systematic reliability and safety assessment with respect to EMUs has long been recognized. In an effort to avoid economic losses and heavy casualties arising from safety violations, a large number of studies have been conducted to combine risk-based reliability analysis into safety control of EMUs. For example, Hanmin Lee, EuijinJoung, et al [18] built the management system in PDM (Product Data management) for failure history data to analyze the reliability of advanced EMU. Joung, E.[14], on the basis of the referenced RAMS standards, presented a system of reliability prediction and relevant demonstration procedure

to apply it to the advanced EMU. As for the traction system, SeoS I, Park C S, Choi S H, et al. [25, 26] offered a procedure that can be utilized to assess and manage in a practical manner the reliability regarding the prototype system of train traction. With the application of reliability block diagram and failure mode effect, an analysis of reliability was carried out after the electric traction system is classified into subsystems. Chateauneuf Aet al. [8] put forward a methodology with the characteristic of consistency reliability to conduct an analysis of traction equipments subjected to fatigue, corrosion, and imperfect maintenance operations with a view to improving their inspection based on a balanced cost and reliability. For the purpose of meeting the challenge that different functions of sub-systems are likely to be activated in different contexts, Wang S, Ji Y, Dong W.[30] devised a new model of reliability analysis which is based on stochastic automata for the traction system of high-speed train. With the theory of stress–intensity distribution interference in mind, and the use of the mode of advanced first order and second moment, Li Cet al. [19] formulated a reliability model to estimate the reliability of an EMU traction system.

An increasing number of recent studies have emphasized on the estimate of reliability with regard to EMUs with the use of Bayesian networks (BN), a widely applied system to conduct uncertain knowledge representation and reasoning. A comparison has been made between the modeling and analysis of fault-trees[23,4], reliability block diagrams[29] and BNs, which has established that they possess a significant advantage over the traditional frameworks. Bayesian network models designed for reliability evaluation can be obtained through the conversion of the traditional ones [6]. Bobbio et al. [4] presented an algorithm of mapping a fault tree with general gates model into an equivalent Bayesian network representationvand Kim [21] presented a general method to convert a reliability block diagram to a Bayesian network. Based on the theory of BN, reference [35] makes the reliability assessment for high-speed train bearing under the extreme sample size which solves the reliability life of the bearing; Reference [34] assesses the reliability of key structure of C70 gondola car which follows Gaussian distributions for extreme sample size. Dorociak, R.[9] presented a method to analyze in a probabilistic manner the reliability regarding an innovative autonomous railway vehicle. It renders a support to the modeling of the failure propagation within the specification of complex systems. When the failure propagation is translated into a Bayesian network, a sophisticated probability analysis is made possible. Guo J, Wilson A G. [12] proposed a Bayesian approach to make an assessment of the reliability of multi-component systems, which facilitates us, with the use of the multilevel information available, to evaluate the system, the subsystem, and the component reliability. We can safely confirm that the research into railway vehicle reliability is still in BN stage.

However, in conventional BN-based analysis, it is a static model which stands for a joint probability distribution at a time interval or a fixed point. Meanwhile, the dependency among variables is not submitted[20,21] in the construction of a BN model. [32, 33]. Yet Dynamic Bayesian Networks (DBNs) are enduring extension of BNs, which make it convenient for us to acquire explicit modeling of temporal dependencies. DBNs render us an unique technique to model time-dependent changes in an intuitive way by means of a robust probabilistic framework [16]. And, the learning and reasoning engine of DBN makes it a possibility for complex interactions among the components of EMUs to be taken into consideration with regard to reliability assessment. Many studies, with the use of DBN to estimate system reliability, have been put forward, while Boudali H. et al. [5] presented that, through the transformation of Dynamic Fault Trees (DFT) into DBN, the integration of the dynamic aspect shall be acquired. As for Portinale et al. [24], a software called Reliability Analysis with Dynamic Bayesian Networks (RADYBAN) has been employed, it supports an automatic translation from DFT into a DBN and presents a method of reliability modeling. Weber and Jouffe [31] had a methodology employed to facilitate a developing dynamic object oriented Bayesian networks to formalize complex and dynamic models, with the model structure deduced from the malfunctioning (knowledge represented through FMECA method) and functional analysis (knowledge formalized by SADT method).

One of the major limitations of the DBN framework, however, is the complexity of the system that can be tractably modeled as a DBN [22]. Current tools, which are based on the assumption of a pre-built DBN, have some disadvantages, including a lack of effective modeling power in depicting both functional and temporal dependencies between components. In this paper, a new modeling approach to DBN generation is submitted, which can be applied to the system made up of certain components and different types of flows propagating through them, with each component possessing Conditional Probability Table (CPT) for a description of the relations between input–output flows. The Component-based CPT (Conditional Probability Table) and Time-dependent CPT are used to describe functional dependencies and temporal dependencies respectively. As the complexity and size of the system cannot be modeled in a tractable way as a DBN, a Breadth-First-Search (BFS) algorithm is introduced for the construction of the DBN model in an automated manner.

The paper has the structure of organization as follows. Section 2 offers a general overview of DBN and its inference scheme. In Section 3, Basic concepts of a dynamic system model and the component-based CPTs(Conditional Probability Tables) are introduced. Based on these concepts and a trace- back algorithm, an automated DBN generation procedure is proposed in Section 4. In Section 5, an application of the proposed approach to reliability analysis is conducted in a traction drive system of the case study of EMUs, with the conclusions reached in Section 6.

2. An overview of DBN

Bayesian Networks (BNs) are in the form of directed acyclic graphs for uncertainty reasoning, with the nodes representing variables and links defining probabilistic dependences between variables. The CPTs connected to the nodes determine the extent to which the linked nodes depend on each other. On the basis of the conditional independency theorem, BN carries out a factorization of the joint probability distribution of a set of random variables $\{X_1, X_2, \cdots X_n\}$ with local dependencies in mind. In this aspect, the distribution mentioned above can be allowed to be decomposed as what is derived from the probabilities of the nodes given their immediate parents:

$$P[X_1, X_2, \cdots X_n] = \prod_{i=1}^{n} P[X_i \mid Pa(X_i)] \tag{1}$$

Where $Pa(X_i)$ is the parent node of variable X_i. The computation is based on the probabilities of the parent's states and the CPT. For instance, let us consider two nodes X_1 and X_2 with two states (S_1 and S_2) each structuring the BN. This CPT is defined as a matrix:

$$P[X_2 \mid Pa(X_2)] = \begin{bmatrix} p(X_2 = S_1 \mid X_1 = S_1) & p(X_2 = S_1 \mid X_1 = S_2) \\ p(X_2 = S_2 \mid X_1 = S_1) & p(X_2 = S_2 \mid X_1 = S_2) \end{bmatrix} \tag{2}$$

By means of relevant temporal dependencies that capture the dynamic behavior of the domain variables between representations of the static network at different times, the static BN can be extended to get a DBN model. Early work in regard to the application of BNs to dynamic domains [1, 2, 10, 14] has rendered us formalisms of DBN, with two types of approaches distinguished in the representations of a

dynamic Bayesian network, i.e., instant-based (time-sliced) type and interval-based (event-based) type [5]. The former involves discretizing the time line and associating a node to every time instant. Basically, the models have been acquired with the generation of a BN for a specific time instant, with the same structure repeated for every time instant over the time range of interest. Example includes Temporal Bayesian Networks (TBN) [15], Modifiable Temporal Belief Networks (MTBN) [1]and Dynamic Object Oriented BN (DOOBN)[31]. In Ref. [31], P. Webber et al. applied a 2-time-slice DBN to model temporal dependencies , with the model structure deduced from the functional analysis. Another representation of DBN is event-based approach. As for the latter, the time line is sliced into a finite number of time intervals, with just one BN generated, and each node possessing a finite number of states equalizing to that of time intervals. The involved examples are Temporal Nodes Bayesian Networks (TNBN) [2], Net of Irreversible Events in Discrete Time (NIEDT)[10] and Discrete-Time Bayesian network (DTBN)[23], where a node stands for an event, with a certain outcome to take place at a certain time interval. According to the fact that our DBN model features the fixed structure that can be repeated for every time instant, this paper adopts the time-sliced type.

The DBNs allow us to taking time into consideration, with the definition of different nodes to stand for the variables at different time slices. The joint distribution of probability regarding a set of randomly variables at time $t+\Delta t$ can be decomposed as what is derived from the probabilities of the nodes given their immediate parents:

$$P\left[X_1^{t+\Delta t}, X_2^{t+\Delta t}, \cdots X_n^{t+\Delta t}\right] = \prod_{i=1}^{n} P\left[X_i^{t+\Delta t} \mid X_i^t, Pa(X_i^t), Pa(X_i^{t+\Delta t})\right] \quad (3)$$

where $X_i^{t+\Delta t}$ and X_i^t are the copies of X_i in two consecutive time slices with a time interval of Δt, $Pa(X_i^t)$ and $Pa(X_i^{t+\Delta t})$ are the parent sets of at the time slices t and $t+\Delta t$ respectively. Depending on the dynamic transition and physical features of the stochastic process of interest and thus the conditional inter- dependencies that need to be modeled, either aforementioned parent set could be empty.

Defining these impacts as transition probabilities between the states of the variable at time step $t+\Delta t$ and those at time slice t leads to the definition of CPTs that are relative to inter-time slices. With this model, the future slice $t+\Delta t$ is conditionally independent of the past given the present t, which means that the CPT respects the Markov properties. This CPT is defined as:

$$P\left(X^{t+\Delta t} \mid X^t\right) = \prod_{i=1}^{N} P\left(X_i^t \mid Pa(X_i^t)\right) \quad (4)$$

where X_i^t is the i th node at time t and $Pa(X_i^t)$ are the parents of X_i^t in the graph. The nodes in the first slice of a DBN do not have any parameters associated with them, but each node in the second slice of the DBN has a conditional probability table (CPT) for discrete variables, which defines $P\left(X_i^t \mid Pa(X_i^t)\right)$ for all $t \geq 1$. Several inference methods for a DBN can be used, i.e., forwards-backwards algorithm, unrolled junction tree, and the frontier algorithm. For the evaluation of the DBN presented in this article, a Netica procedure[7] based on the **junction tree** is used.

3. Component-based CPT

This modeling method involves a component-based approach, i.e., any system is modeled as a group of interconnected components. Different classifications of energy, information or materials are transferred in the form of flows by means of the connections between the components, with each one possessing some input or output relationships that quantified through the association of a conditional probability table. Generally speaking, components in the system serve as the building blocks and flows make their movements among the components.

3.1. Basic concepts

In order to define Component-based CPT for DBN formally, we, first, capture the breakdown structure of the system into physical or functional components; second, identify the input and output flows including types and finally model the system as networked components with input and output flows.

3.1.1. Components

A component is defined as any of the elementary unit of a system, including all the electrical and mechanical devices. They are in the form of either active components such as coolers and pumps, or passive components such as wires and pipes. Symbolized as a simple circle, each component has a label, with some arrows pointing outward or inward, acting as its output or input paths in a respective manner. CPT, the vital part of component models, defines how output values are formed on the basis of input values. Here three kinds of components are defined: flow-intervened components, flow-collaborated components and flow-dependent components. Each one can take on several states or failed modes, and as far as each state, there is a definition of time-dependent CPT. Take the pressure control as an example. It involves several failed modes identified from previous experience and expert judgment. An expert can predict at least the following software malfunctions, i.e., the low-stuck, high-stuck or oscillating control signal.

3.1.1. Flows

A system can be considered as a networked topology structure of components along flow paths. A flow is generally defined as any energy, information or materials propagating from one component to another. Various discrete scales have been put in place in the previous work in engineering design to identify design dependencies between components in respect to the flows of material, energy or information among functional components of systems during their concept development[28]. Energy, matter and information are considered basic concepts in any design problem. It is the flow of these three concepts that concerns designers. Flows are equipped with some physical properties vital to system analysis. As regards components, they can have an effect on these properties with flows passing through them. For example, an oil flow has a possession of physical properties like pressure, temperature, concentration and the flow rate. Likewise, their range of variation is expected to get specified.

3.2. Component-based CPT

Component-based CPTs are defined in accordance with the types of basic components tying together the input and output flows. The use of a component-based CPT allows us to obtain the output values for combination of varying input. And thanks to the component-based CPTs, we can acquire a model of the propagation involving several failure modes in the system through the input-output flows. Then, the hypothesis of independence between components made for traditional reliability assessment is not necessary. As a matter of fact, component-based CPTs make it possible for us to compute repercussions of

interdependence components to the system reliability and introduce uncertainty by putting probabilities in place at the interval of value [0, 1].

3.1.1. Flow-collaborated components and its BN formalization

A sample Component-based CPT for a flow-collaborated component is shown in Fig.1(a) The flow-intervened component has the same type of flow input and output. The output serves as a function or functionality condition in respect to the inputs. With the input from another component of the system, the CPT is an internal parameter embedded in the component. For the purpose of reaching the conversion algorithm from flow-intervened component to BN, we have adopted the following convention: Given a generic binary input flow or component CMP, we denote with CMP=1 the component failure and with CMP =0 the component normal. With the usual hypothesis that input flow or component failures are distributed in an exponential way, the probability of the occurrence of the primary event CMP=1 at time t is $\exp(-\lambda \times t)$, where λ acts as the failure rate of input flow or component.

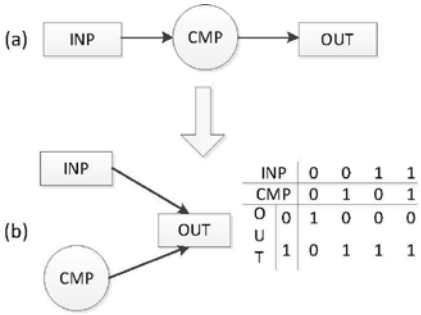


Fig. 1. Component-based CPT for a Flow-collaborated Component

Fig.1(b) demonstrated the conversion of a flow-collaborated component. Input flow(INP)and flow-collaborated component (CMP) are assigned to probabilities in advance(in agreement with the probability of the occurrence of the primary event INP =1 or CMP=1), and output flow is assigned to its CPT. The output fails when flow-collaborated component or input flow falls into a failed state.

3.2.1. Flow-intervened components and its BN formalization

A sample Component-based CPT for a flow-intervened component is shown in Fig.2(a). The flow-intervened component has one type of flow input and another type of flow output. The output is a function or functionality condition of the inputs. With the input coming from another component of the system, the CPT acts as an internal parameter of the component. In order to reach the conversion algorithm from flow-intervened component to BN, we adopt the same convention mentioned above.

Fig.2(b) shows the conversion of a flow-intervened component. Input flow(INP) is assigned prior probabilities (coincident with the probability of occurrence of the primary event INP=1), and flow-intervened component are output flow are assigned CPT. CSCSIT[13] is potentially to help improve the Flow-intervened components' CPT modeling for that it can provide a reliability parametrized component-based modeling structure. In CSCSIT, a component failure occurs when the conceptual stress of the input flow exceeds its conceptual strength. Therefore, the conditional probability P_{int} for the component CMP failure is given by:

$$P_{\text{int}} = \Pr(CSte_{INP} \ \geq \ CStn_{CMP}) \tag{5}$$

Here $CSte_{INP}$ is the conceptual stress of the input flow and $CStn_{CMP}$

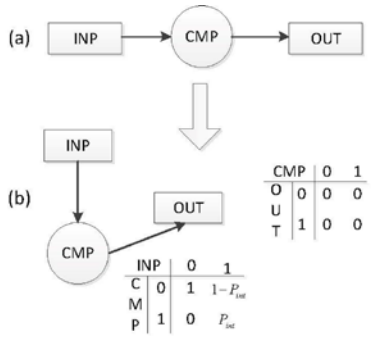


Fig. 2. Component-based CPT for a flow-intervened component

is the conceptual strength for the component CMP. One nice feature of the CSCSIT is that it can represent the uncertainty by conditional probability P_{int} .

3.3. Time-dependent CPT

In the simplest form, as for the component in Section3.2, the component is either in failed or work state. However, there is over one failure mode for a component in the general case. For example, a traction motor may be either overheating (O) or in operation (I). In our work, the failure modes of a component are defined by states, with a description of its ability or inability to output the desired flows. The multi-state of components can be modeled as a DBN formalization.

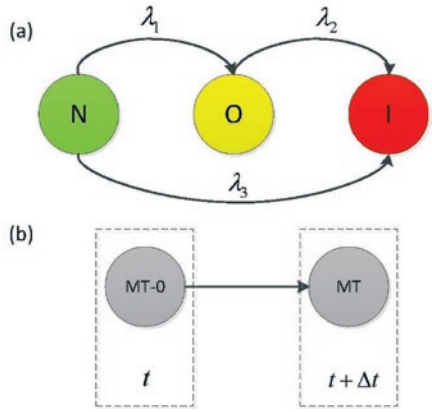


Fig.3 Markov model for state transition modeling of traction motor

Consider a traction motor with three states, which can be envisaged as: normal (N), overheating (O), and in operation (I). An Markov model of traction motor reliability is easy to build, which is shown in Fig.3(a). Then, independent components of the process are modeled by using DBN that is equivalent to an independent MC: Firstly, the traction motor is modeled by a discrete random variable X with states {N, O, I}. Next two nodes are defined to model the random variable at time slices t and $t+\Delta t$: $CMP(t)$ and $CMP(t+\Delta t)$. Kinked by an arc representing the dependency between the component states at time slice t and its states at time slice $t+\Delta t$, these nodes are both rendered a description by the states {N, O, I}. With the assumption of the constant failure rates of normal (N), overheating (O) and in operation (I) denoted as λ_1 , λ_2 and λ_3 , usually estimated by using historical data, the aforementioned transition gets modeled in the corresponding DBN through the conditional probability. In order to get clarified, the Markov model and the equivalent DBN in regard to state transition modeling of the traction motor are demonstrated in Fig.3(b), with T1 at $t+\Delta t$, the time-dependent conditional probabilities, presented in Table1.

Table 1. Time-dependent CPT of the Traction Motor at $t+\Delta t$

$CMP(t)$		N	O	I
$CMP(t+\Delta t)$	N	$\exp(-\lambda_1\times\Delta t)+\exp(-\lambda_3\times\Delta t)-1$	$1-\exp(-\lambda_1\times\Delta t)$	$1-\exp(-\lambda_3\times\Delta t)$
	O	0	$\exp(-\lambda_2\times\Delta t)$	$1-\exp(-\lambda_2\times\Delta t)$
	I	0	0	1

4. Modeling approach

Previous sections described the necessary details and settings of modeling a system. Her we move on to the second phase involving the procedure of the automated DBN construction, where a Breadth-First-Search (BFS) algorithm works on the provided model to generate the Structure of DBN in the system. The depth-first search is used to identify reencountered events, and practise some simplification in the end. The main interest in such an approach to render a reliability modeling from DBN consists in the propagation of the components' failure through the input/output flow of the system. However, the operation to model complex systems makes a methodology a necessity in an effort to get the DBN's structure specified, with four main steps involved in the proposed modeling approach as follows:

1. Create the model of component and input/output flows for the analysis of system.
2. Identify the target event and system boundaries that have an active influence on the propagation of components' failures.
3. Structure a learning of DBN based on BFS.
4. We will endeavor to formalize the DBN from this System representation.

4.1. Step 1: System modeling

With the previous concepts introduced, we are allowed to model a variety of complex technical systems, including both configuration network and CPT description. The former's configuration is simply modeled with appropriate components to get connected to each other in a desired manner. The components regarding CPT will simulate how the system operates in a comprehensive way. As far as the system modeling in DBN construction is concerned, we will have two procedures illustrated in an explicit way in this step.

1) Malfunction knowledge acquisition: Previous resources in regard to malfunctioning knowledge are supposed to be acquired in the first place in an effort to offer a basis for System modeling. The resources can be obtained in two aspects: empirical knowledge from domain experts and the accepted knowledge from existing design and construction standards. The collected knowledge can be expressed in table.2 The components of EMU traction system are listed in the first row and the column in table.2, with their meanings queried to what is listed in Section 5.1 and the connection between components is represented by the flow of energy, matter and information.

2) Configuration model creation: Developed from a direct mapping of the functional model to generic components by using the input and output flows, the configuration model represents the actual design under consideration. Called the component basis and the flow taxonomy, a dictionary of flows and components as introduced in [27,17], is presented in the methodology involved in this paper, where the configuration layout of the system is captured by means of configuration flow graphs (CFGs), with blocks representing system components and lines connecting the blocks the material, energy, or information among the components(to be clarified, with solid lines representing material, dashed lines energy and dash dotted lines signals). An example of CFG is shown in Fig.4.

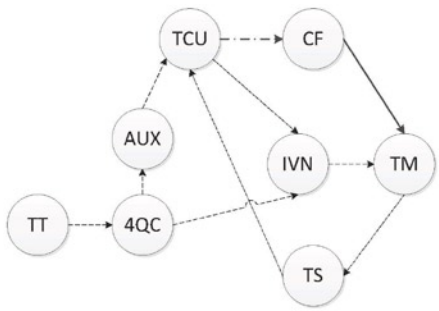


Fig. 4. Configuration flow graphs of EMU traction system

Table 2. Component connection table of EMU traction system

Components	TT	4QC	AUX	TCU	INV	CF	TS	TM
TT		Energy						
4QC			Energy		Energy			
AUX				Energy				
TCU					Information	Information		
INV								Energy
CF								Matter
TS				Information				
TM							Information	

4.2. Step 2: Target event identification

A target event is defined by assigning specific values to some of the system parameters. The target identification process involves two steps:

1) System boundaries identification: we firstly determine the border of configuration model. The hypothetical lines, which are covered by BFS algorithm to determine the lines, are called system boundaries. A component located at system boundary is a starting point for the BFS algorithm.

2) Target event: We can specify any state of the starting point as a target event, including some parameters of a flow, like temperature pressure, flow rate, or pressure. The values can be selected within the range of these parameters. Under the definition of a sample target event, the output of pump p2 is zero or the temperature of flow f1 is high. Then, a DBN synthesis algorithm starts from the point of occurrence of the target event, and relevant components, after which

the system reliability condition can be determined by tracing back all possible paths ending into the target event.

4.3. Step 3: DBN synthesis

A DBN construction is developed in this stage to integrate the configuration of the system in a simultaneous way, with the target event defined by what is completed in the previous step. What is mainly involved in procedure of the DBN construction is a BFS algorithm to acquire all the paths leading into the target event. Depth-first search is one of the basic algorithms regarding graph theory, which is commonly employed to a test in connectivity or compute the shortest paths of underweighted graphs of single source [3]. With the target event is defined and all of the required CPT settings completed, the BFS algorithm is supposed to start from the point of occurrence of the target event, and examine the CPT of the components to acquire all of their potential causes. The algorithm switches from one component to another in a reve rse direction of the flows. The synthesis of a DBN construction involves a determination of the system boundaries, network structure, as well as the Markov chain process, during which process three procedures are illustrated explicitly as follows:

1) Component added (BFS tree): Starting from the target event, the frontier of BFS makes an outward expansion with each step, visiting all of the same-depth components before visiting another at the next depth. Within a step of such top-down approach, each component checks all of its neighbors to see if there are overlooked ones, with its CPT searching to locate the rows with the output value of interest. In case of a CPT with one component leading to that output value, the previously unvisited component will be added to the frontier and marked as visited by setting its parent variable. The procedure is continued until each component at the system boundary is touched. This algorithm yields a BFS tree and The pseudo code is detailed as Table.3.

Table 3. Pseudo code of BFS

```
breadth-first-search (components, target event)
frontier ⟵ { target event }
next⟵ {}
parents⟵ [−1,−1, . . . ,−1]
while frontier= {} do
for c ∈ frontier do
for n ∈ neighbors[c] do
if parents[n] = −1 then
parents[n]⟵ c
next⟵ next ∪ {n}
end if
end for
end for
frontier⟵ next
next⟵ {}
end while
return tree
```

2) DBN Structure learning:

Structure learning aims at figuring out a proper directed acyclic graph (DAG), and confirming the failure mechanism among nodes. In accordance with the establishment of BFS tree, explicit DAG and failure mechanism among the components can then be revealed by the input and output flows. The definition of the failure mechanism is formalized at the level of the system, while the description of the failure mode is made at the level of component. In line with this functioning, the malfunctioning of the system is induced with a consideration of the normal and abnormal states of the components, with every component in the real-world situation represented by a Bayesian node. When a component is considered in the BFS algorithm, the BN formalization of its CPT is created, with the rows involving the output value of interest set as its inputs. Each of these inputs will be checked for need to further expand. However, with this multi-state components considered, the DBN formalization of its CPT is created and two nodes are defined to model the variable in a random way at time slices t and $t+\Delta t$. An example of traction motor(TM) is shown in Fig. 5.

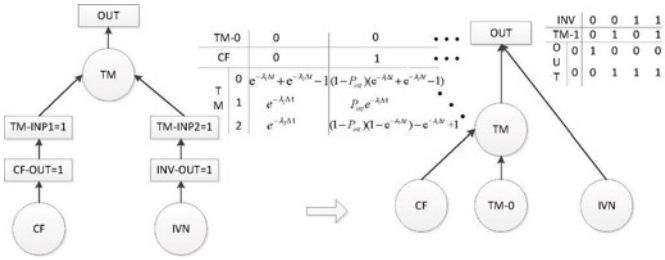


Fig. 5. Structure Learning of Traction Motor(TM)

In Fig. 5, there are CF (cooling fan), TM Flow-collaborated, INV (traction inverter) and TMF low-intervened components. TM itself is a Time-dependent component. Based on the generation principle of CPT, Fig. 5(a) can be transformed into Fig. 5(b) by structure learning.

5. Example

5.1. DBN model of EMU traction system

The CRH5 high-speed EMU is designed for a speed of 250 km/h, which consists of two symmetrical traction units (Mc(1), M2(2), TP(3)and M2(4) comprise tractions of unit 1; T2(5), TPB(6), MH(7) and Mc(8) comprise traction unit 2).Modern mechatronics technology and its new features have led to a continuously improvement of the construction of the traction unit structures. A traction unit mainly includes traction transformer(TT), auxiliary inverter(AUX), traction inverter(INV), traction control unit(TCU), four-quardant rectifier(4QC), cooling fan(CF), traction motor(TM) and its temperature sensor(TS). The traction unit can be highly complex due to the systematic use of new technologies and be functional dependency due to the interactions between system functions with the characteristics of a variety of structures.

In traction unit 1, four-quardant rectifier obtains power through the traction transformer at the bottom of the vehicle, with the power transferred to the traction inverter and the auxiliary inverter. Traction control unit gets power through the auxiliary inverter and control the air volume of fan following the signals collected by temperature sensor, Traction inverter is connected to the traction motors through the terminal blocks of traction inverter. The traction motors have a cooling fan and a temperature sensor in order to monitor and reduce the motor temperature. Mechanical energy from the output of the traction motor is delivered to the wheels through the OUT1. Fig. 6(a) shows the configuration model of the traction system of CRH5. According to the DBN modeling approach proposed in Section 4, we have chosen

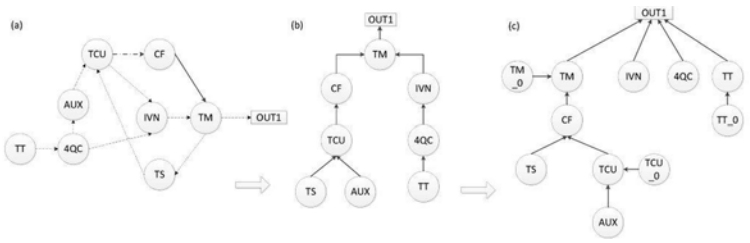


Fig. 6. DBN Model of EMU Traction System

energy output of traction motors OUT1 as a target event and a certain BFS tree is shown in Fig. 6(b).Then, the structure learning of DBN is based on the DBN formalization of components' CPTs, and finally, we obtained the DBN shown as Fig. 6(c).

5.2. Reliability evaluation

Within the failure statistic data of the traction system of CRH5 which is running in Beijing- Harbin high-speed railway, we have conducted a calculation of the reliability indexes of the CRH5 [4,6,21], with the time range set from September 15, 2011 to June 20, 2015 respectively. With such an immense amount of data, we have just presented the results with respect to the reliability indexes of the components mounted on the traction of unit 1 which is running in Beijing-Harbin railway. The results are presented in Table.4. In Table.4, where MDBF is the abbreviation of Mean Distance Between Failures.

Table.4 Reliability Indexes of the Components

Component	Number of failure	Average failure Rate(Time/1E5km)	MDBF(1E5km)
TM	89	0.003415	292.8258
TCU	113	0.005621	177.9043
TT	77	0.002906	344.1156
INV	85	0.003205	312.0125
4QC	75	0.002655	376.6478
TS	121	0.006528	153.1863
AUX	80	0.003137	318.7759
CF	31	0.001312	762.1951

The reliability evaluation phase of the traction system is established with the application of Netica software shown in Fig.7. Take the traction system as an example, parent nodes TM_0, TS, AUX,TCU_0, TT_0, INV and 4QC at the component state layer stand for the states of components, excluding the faults of components. OUT, the unique child node represents the state of the traction system. Child node OUT has two states, i.e., normal one and failed one, and the probability that the normal one of OUT represents the value of the system reliability. For the purpose of modeling the temporal evolution of a system, we have selected two time slices for multi-states components, e.g., TM represents the current time step of traction motor, and TM_0 the previous time step. The time interval Δt could be 1E5 km or 1E5 h. A wealth of time slices is in agreement with a smaller the value of Δt, hence rendering Netica

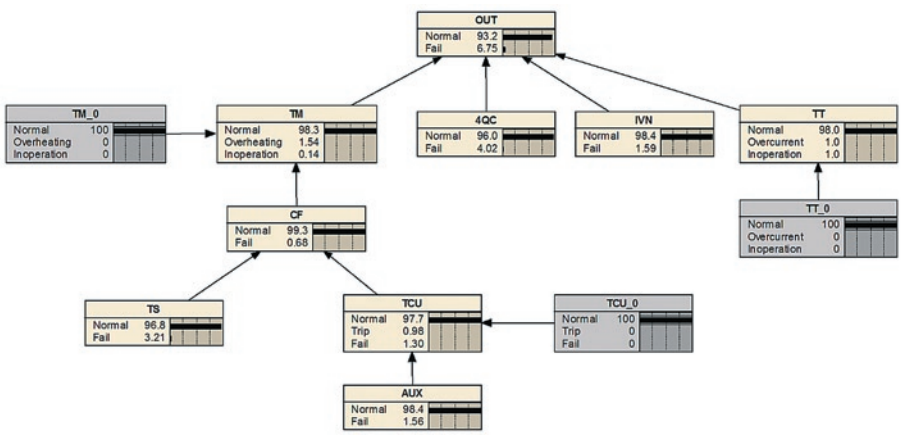


Fig. 7. The Reliability Evaluation Phase of the Traction System

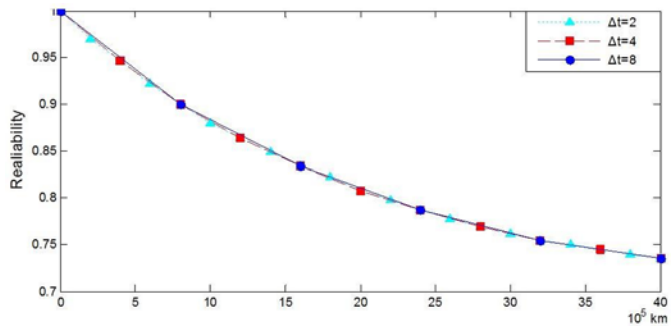


Fig. 8. Reliability of the Traction System without Components' Failure

a longer running time. Here the DBNs are extended 1E5 km, and the DBNs structure of what is extended within the 40E5 km is presented in Fig.8.

When no component failure occurs, time interval, Δt, the system reliability is rarely affected for the reason that the extension of DBNs and the conditional probability of time slices involving each component are rooted on the exponential distribution of components. As we have expected, with the increase in time, the reliability of the traction system decreases in a corresponding way. According to the DBN model of the traction system and what is obtained from the reliability indexes of the components of traction system, we can calculate that: In the 40E5km, the system still possess a reliability of 73.326%, with an indication of t sufficient secure on the part of the traction system.

When a certain component is abnormal, a calculation and plot of the reliability values shall be carried out. As is shown in Fig.8 (a), there comes a rapid decrease of reliability before the overheating of

Table. 5 Two Case of Reliability Evaluation

Component	TM	TCU	TT	INV	4QC	TS	AUX	CF
Case1	Overheating	Normal	Normal	Normal	Normal	Normal	Normal	Normal
Case2	Normal	Normal	Normal	Normal	Normal	Normal	Fail	Normal

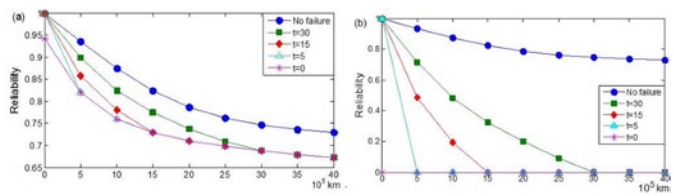


Fig. 9. Reliability of the traction system when a certain component is abnormal

traction motor is detected, and once the traction motor goes in operation, the reliability of traction system decreases to 0 with the immediate overheating of traction motor after the system is started, rendering the motor a fatal weakness of the traction system and making it a necessity the improvement of the reliability of the traction motor value to the fullest degree.

As is illustrated in Fig.8 (b), there comes the fault of auxiliary inverter immediately after the system is started, with the reliability decreasing to a lower level lower when the no fault occurs. In the 40E5km, the system still possesses a reliability of 68.28%. With the occurrence of the fault at a certain intermediate time, the reliability decreases to the value when the fault occurs at $t = 0$. This arises from the fact that the auxiliary inverter is a flow-intervened component,

and its failure has little effect on the normal operation of the traction system, but with a decrease of the reliability regarding the entire system.

6. Conclusion

DBN serves as a powerful tool for knowledge representation and reasoning in a complex mechatronic system. A new system of modeling approach for DBN generation is presented in this paper, with the introduction of a component-based configuration model made up of some components and different types of flows propagating through them and a Breadth-First-Search (BFS) algorithm for the automated construction of the DBN model. The Configuration model comprises some components and different types of flows propagating through them, with each component possessing a CPT description of its input–output flows relations. As the size and complexity of the system cannot be tractably modeled as a BN, a Breadth-First-Search (BFS) algorithm is introduced for automated construction of the DBN model. Given that traditional DBN framework can not be tractably modeled, our method offers a good description of functional and temporal dependencies between components, which turns out to be a satisfying solution with regard to the modeling of complex systems.

We have demonstrated in this paper the application of the proposed approach to reliability assessment of traction drive system regarding high-speed EMUs, with the reliabilities evaluated by means of netica, taking into account two important features of dependency between components and multi-state components. An application of the proposed Bayesian network models facilitates the evaluation of the reliability of traction drive system at any given time. Results have established that the DBN-based approach can perform in a more accurate way than the traditional static one in regard to modeling the evolution of the probabilistic dependencies within a complex system over time.

It may be concluded that, with the use of DBN in the reliability assessment of complex mechatronic systems, we can not only avoid problems such as the failures' dependencies and the multi-state elements, a common thing in traditional static approach, but also help analysts to conduct probability updating, which is of great significance in the real-time monitoring, and evaluation of mechatronic systems. Future work will place an emphasis on the investigation of a real-time intelligent reliability evaluation software with the application of automatic data acquisition and the proposed DBN modeling approach for an entire system, such as a EMU system equipped with thousands of flow-intervened or flow-collaborated components.

Acknowledgment

The authors gratefully acknowledge the support provided by the research project (Grant No. RCS 2016ZZ002) of State Key Laboratory of Rail Traffic Control and Safety (RCS) in China.

References

1. Aliferis C F, Cooper G F. A Structurally and Temporally Extended Bayesian Belief Network Model: Definitions, Properties, and Modeling Techniques. Eprint Arxiv, 2013.
2. Arroyo-Figueroa G, Sucar L E. A temporal Bayesian network for diagnosis and prediction. Fifteenth Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann Publishers Inc. 1999:13-20.
3. Beamer S, Asanović K, Patterson D. Direction-optimizing breadth-first search. Scientific Programming 2013; 21(3-4): 137-148, https://doi.org/10.1155/2013/702694.
4. Bobbio A, Portinale L, Minichino M, et al. Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. Reliability Engineering & System Safety 2001; 71(3): 249-260, https://doi.org/10.1016/S0951-8320(00)00077-6.
5. Boudali H, Dugan J B. A discrete-time Bayesian network reliability modeling and analysis framework. Reliability Engineering & System Safety 2005; 87(3): 337-349, https://doi.org/10.1016/j.ress.2004.06.004.
6. Cai B, Liu Y, Liu Z, et al. 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. Cai B, Liu Y, Ma Y, et al. Real-time reliability evaluation methodology based on dynamic Bayesian networks: A case study of a subsea pipe ram BOP system. Isa Transactions 2015; 58: 595-604, https://doi.org/10.1016/j.isatra.2015.06.011.
8. Chateauneuf A, Cocheteux F, Deffarges F, et al. Reliability analysis of screwed connections in high-speed trains, considering fatigue, corrosion, and imperfect maintenance operations. Proceedings of the Institution of Mechanical Engineers Part O Journal of Risk & Reliability 2011; 225(3): 293-306, https://doi.org/10.1177/1748006X11402738.
9. Dorociak R. Early probabilistic reliability analysis of mechatronic systems. Reliability & Maintainability Symposium. IEEE, 2012:1-6.
10. Gala'n SF, Dı'ez FJ. Modeling dynamic causal interactions with Bayesian networks: temporal noisy gates. CaNew', the 2nd International Workshop on Causal Networks held in conjunction with ECAI 2000, Berlin, Germany, August 2000; 1–5.
11. Guangzu S. Efforts to create a new situation of scientific development of railway, better services for economic and social development and the people. Railway Economics Research 2012; 1: 1-26.
12. Guo J, Wilson A G. Bayesian methods for estimating system reliability using heterogeneous multilevel information. Technometrics 2013, 55(4): 461-472, https://doi.org/10.1080/00401706.2013.804441.
13. Huang Z, Jin Y. Extension of stress and strength interference theory for conceptual design-for-reliability. Journal of Mechanical Design 2009; 131(7): 1-11.
14. Joung E, Kim G, Lee J, et al. Reliability Analysis of New Type Electric Multiple Unit. Computer Applications for Modeling, Simulation, and Automobile 2012: 213-220,https://doi.org/10.1007/978-3-642-35248-5_30.
15. Kanazawa K. Reasoning about time and probability. Brown University, 1994.
16. Khakzad N. Application of dynamic Bayesian network to risk analysis of domino effects in chemical infrastructures. Reliability Engineering & System Safety 2015; 138: 263-272, https://doi.org/10.1016/j.ress.2015.02.007.
17. Kurtoglu T, Tumer I Y. A Graph-Based Fault Identification and Propagation Framework for Functional Design of Complex Systems. Journal of Mechanical Design 2008; 130(5): 680-682, https://doi.org/10.1115/1.2885181.
18. Lee H, Joung E, Kim G, et al. A study on management system for reliability analysis in advanced EMU. Vehicle Power and Propulsion Conference (VPPC) 2012 IEEE. IEEE, 2012: 1266-1269.
19. Li C, Qiao C, Zhang Y, et al. Reliability optimization design of connecting rod of locomotive traction equipment. Archive Proceedings of the Institution of Mechanical Engineers Part C Journal of Mechanical Engineering Science 1989-1996; 203-210.

20. Liu J, Li S, Jiang Y, et al. Reliability evaluating for traction drive system of high-speed electrical multiple units. Transportation Electrification Conference and Expo (ITEC), 2013 IEEE. IEEE, 2013: 1-6.
21. Man C K. Reliability block diagram with general gates and its application to system reliability analysis. Annals of Nuclear Energy 2011; 38(11): 2456-2461, https://doi.org/10.1016/j.anucene.2011.07.013.
22. Marquez D, Neil M, Fenton N. Improved reliability modeling using Bayesian networks and dynamic discretization. Reliability Engineering & System Safety 2010; 95(4): 412-425, https://doi.org/10.1016/j.ress.2009.11.012.
23. Portinale L, Bobbio A. Bayesian Networks for Dependability Analysis: an Application to Digital Control Reliability. Computer Science 2013: 551—558.
24. Portinale L, Raiteri D C, Montani S. Supporting reliability engineers in exploiting the power of Dynamic Bayesian Networks. International Journal of Approximate Reasoning 2010; 51(2): 179-195, https://doi.org/10.1016/j.ijar.2009.05.009.
25. Seo S I, Park C S, Choi S H, et al. Reliability management and assessment for the electric traction system on the Korea High-Speed Train. Proceedings of the Institution of Mechanical Engineers Part F Journal of Rail & Rapid Transit 2010; 224(3): 179-188, https://doi.org/10.1243/09544097JRRT297.
26. Seo S I, Park C S, Han Y J, et al. Reliability assessment of traction system of Korean high speed train. Journal of the Korean society for railway 2005; 8(5): 434-438.
27. Sierla S, Tumer I, Papakonstantinou N, et al. Early integration of safety to the mechatronic system design process by the functional failure identification and propagation framework. Mechatronics 2012; 22(2): 137-151, https://doi.org/10.1016/j.mechatronics.2012.01.003.
28. Stone R B, Wood K L. Development of a functional basis for design. Journal of Mechanical Design 2000; 122(4): 359-370, https://doi.org/10.1115/1.1289637
29. Torres-Toledano, Sucar L E. Bayesian Networks for Reliability Analysis of Complex Systems. Lecture Notes in Computer Science 1997; 1484(1484): 465-465.
30. Wang S, Ji Y, Dong W. Reliability Analysis for High-Speed Train Control System by Means of Stochastic Automata. Advanced Science Letters 2012; 6(1): 619-624, https://doi.org/10.1166/asl.2012.2248.
31. Weber P, Jouffe L. Complex system reliability modelling with dynamic object oriented Bayesian networks (DOOBN). Reliability Engineering & System Safety 2006; 91(2): 149-162, https://doi.org/10.1016/j.ress.2005.03.006.
32. Weber P, Medina-Oliva G, Simon C, et al. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. Engineering Applications of Artificial Intelligence 2012; 25(4): 671-682, https://doi.org/10.1016/j.engappai.2010.06.002.
33. Wu X, Liu H, Zhang L, et al. A dynamic Bayesian network based approach to safety decision support in tunnel construction. Reliability Engineering & System Safety 2015; 134: 157-168, https://doi.org/10.1016/j.ress.2014.10.021.
34. Y. Tian, Research on extreme small sample reliability evaluation method and its application on center sill and body bolster of C70 gondola car, Beijing: Beijing Jiaotong University, 2008.
35. Zhu D, Liu H. Reliability evaluation of high-speed train bearing with minimum sample. Journal of Central South University, 2013.

Yanhui WANG
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University 100044, China
School of Traffic and Transportation
Beijing Jiaotong University, Beijing 100044, China

Lifeng BI
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University 100044, China
Jinan Locomotive Depot
Jinan Railway Bureau 250001, China

Shujun WANG
Shuai LIN
Wanxiao XIANG
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University 100044, China
School of Traffic and Transportation
Beijing Jiaotong University, Beijing 100044, China

E-mails: wangyanhui@bjtu.edu.cn, 1028314098@qq.com, 14120954@ bjtu.edu.cn, 13114253@ bjtu.edu.cn, 15120778@bjtu.edu.cn

Article citation info:
RAPOSO H, FARINHA JT, FERREIRA L, GALAR D. An integrated econometric model for bus replacement and determination of reserve fleet size based on predictive maintenance. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 358–368, http://dx.doi.org/10.17531/ein.2017.3.6.

Hugo RAPOSO
José Torres FARINHA
Luís FERREIRA
Diego GALAR

AN INTEGRATED ECONOMETRIC MODEL FOR BUS REPLACEMENT AND DETERMINATION OF RESERVE FLEET SIZE BASED ON PREDICTIVE MAINTENANCE

ZINTEGROWANY EKONOMETRYCZNY MODEL DO MODELOWANIA WYMIANY TABORU AUTOBUSOWEGO ORAZ OKREŚLANIA WIELKOŚCI FLOTY REZERWOWEJ W OPARCIU O KONSERWACJĘ PREDYKCYJNĄ

Maintenance policies influence equipment availability and, thus, they affect a company's capacity for productivity and competitiveness. It is important to optimize the Life Cycle Cost (LCC) of assets, in this case, passenger bus fleets. The paper presents a predictive condition monitoring maintenance approach based on engine oil analysis, to assess the potential impact of this variable on the availability of buses. The approach has implications on maintenance costs during the life of a bus and, consequently, on the determination of the best time for bus replacement. The paper provides an overview of economic replacement models through a global model, with an emphasis on availability and its dependence on maintenance and maintenance costs. These factors help to determine the size of the reserve fleet and guarantee availability.

***Keywords**: condition monitoring, LCC, replacement, reserve fleet.*

Polityka konserwacji wpływa na gotowość sprzętu, a tym samym na wydajność i konkurencyjność przedsiębiorstwa. Ważne jest optymalizowanie kosztów cyklu życia (LCC) aktywów, w tym przypadku taboru autobusowego. W artykule przedstawiono metodę utrzymania ruchu polegającą na predykcyjnym monitorowaniu stanu w oparciu o analizę oleju silnikowego w celu oceny potencjalnego wpływu tej zmiennej na gotowość autobusów. Podejście to ma praktyczne konsekwencje jeśli chodzi o koszty utrzymania w trakcie eksploatacji autobusu, a także pozwala na ustalenie najlepszego czasu na wymianę pojazdów taboru. W pracy przedstawiono przegląd ekonomicznych modeli wymiany oraz opracowano model globalny integrujący te modele, ze szczególnym uwzględnieniem gotowości oraz jej zależności od konserwacji oraz kosztów utrzymania ruchu. Czynniki te pomagają określić wielkość floty rezerwowej i zapewnić gotowość taboru.

***Słowa kluczowe**: monitorowanie stanu, koszty cyklu życia, wymiana, flota rezerwowa.*

1. Introduction

A reserve fleet is defined as the number of vehicles ready to deliver the function for which they were designed, that is, not immobilized by breakdown or planned maintenance, with maximum availability. In transport sector companies, the efficient use of physical assets is linked to a well-structured policy of evaluation and fleet replacement. National and international road transport companies have a wide range of suggested ratios of reserve fleet to total fleet from which to choose. The recommended size of a fleet reserve specified by the US FTA (Federal Transit Administration), in Circular C 1A 9030 1987 Appendix A, is 20% of the total number of vehicles [47].

In the passenger transport sector, the determination of the optimum time for bus replacement is related to the efficient use of assets and the company's global or total costs. The company needs to know the right time to replace a bus to reduce its total costs, while still guaranteeing the availability and quality of service and ensuring customer satisfaction. The main objective of this paper is to define a methodology to determine the best cost at the best time to replace a bus.

The value of money is directly linked to time, because the later an asset is withdrawn from service, the greater the action of external agents on it and the greater the influence of macroeconomic factors (i.e., the inflation rate) on its value. This obviously has an effect on transportation costs.

The paper discusses the relations between some technical maintenance Key Performance Indicators (KPI's) [42], specifically Mean Time to Repair (MTTR), Mean Time Between Failure (MTBF), Availability (A), and the dimension of the reserve fleet. The Return On Investment (ROI), a financial indicator to assess the equipment's financial performance, is used as a "bridge" between the maintenance and the economic fields, showing us when the equipment starts creating profit or loss to the company. It looks at models to determine the influence of these variables on the withdrawal time and size of the reserve fleet and uses oil analysis as an example of how condition monitoring may influence the availability of the whole bus fleet and the size of the reserve fleet. The paper gives a comprehensive summary of the discussed methodologies, emphasizing the immobilization time caused by maintenance; using a global model, it demonstrates its relevance to the dimensioning of the reserve fleet and the withdrawal time.

2. State of the art

Condition Monitoring (CM) represents an approach to preventive maintenance based on knowledge of the equipment's condition determined by monitoring one or several parameters that permit us to

evaluate the equipment's health [12]. The maintenance of a passenger bus is a strategic activity to maximize its life cycle, involving a combination of management, technical and economic actions to achieve high availability at reasonable costs [3, 6, 7, 11, 31, 32].

The Life Cycle Cost (LCC) of an asset represents the sum of all capital spent to support it from design and manufacturing, through its operation until the end of its life, (CAPEx+OPEx - Capital Expenditure+Operational Expenditure), [6]. The LCC of an asset can be significantly higher than the value of the initial investment, and is usually defined at the design phase [7]. Bescherer [11] says as much as 70-90% of the total LCC costs are defined at this stage. To this, Aoudia *et al*. [3] add that poor maintenance management contributes to a significant increase in LCC.

The initial investment cost is often the only criterion in purchase decisions, notwithstanding the benefits of a LCC approach. Possible reasons include a lack of knowledge of the existing standards or formal guidelines and the absence of reliable past data. There are few cross-case studies in the field of life cycle costing, and most are limited to a single industry, [32]. Korpi and Ala-Risku [31] give an overview of the use of LCC and the feasibility of implementation, through a summary of work on its application.

Simply stated, LCC analysis predicts the future. Several methods can be used to estimate future costs, as, for example, Activity-Based Costing (ABC) [20, 21]. Certain standards, such as those specified in [8, 9], support the use of LCC analysis. The rules on asset management given in PAS 55 [28], and in ISO 5500X [8] are good guidelines for physical asset management and can be applied in any sector.

The above points are well known, but there is a lack of systematic study in this area. We need new management models to improve equipment productivity and quality of service, with aspects like environmental sustainability, quality management standards, security, maintenance and energy included in the models [22]. Many companies keep equipment in operation, even when this is no longer economically viable, simply because they do not consider their entire economic cycle [22]. This has implications for many areas, including the size of the reserve fleet.

According to William *et al.* [52], traditional production systems are built on the principle of the economy of scale. The authors illustrate an equipment replacement problem in the context of Lean Thinking, showing the relevance of econometric models. Jennifer and Joseph [29] refer to technological change as a motivator for equipment replacement; they say technology develops continuously according to a well-defined function. Natalia and Yuri [38] combine discrete and continuous models in time to show that the replacement time for equipment decreases when the technology is more advanced.

According to Assaf [5], "the evaluation of an asset is established by the cash flows expected future benefits referred to the present value by a discount rate that reflects the risk of the decision". Consequently, methods considering the value of money over time are the most suitable to use in replacement decisions. In the view of Casarotto [16], the Annual Cost Uniform Equivalent approach is suitable for the analysis of operational activities of a company with investments that can be repeated. The standardization of investment based on annual equivalent values facilitates the analysis required for decision-making. With this method, it is possible to determine which year has the lowest equivalent annual cost; this, in turn, indicates the best technical replacement period [16]. The calculation of the equivalent annual cost is based on the Capital Recovery Factor. It is possible to compare two or more investment opportunities to determine the best time for equipment replacement, taking into account information such as: acquisition value; maintenance cost; resale value or residual value at the end of each year; operating costs; the cost of capital or the attractive minimum rate [49].

To determine equipment economic life with the objective of finding the most rational replacement time, four situations are applicable [36]:

i When the asset is already unsuitable for work;

ii When the asset has reached its lifespan;

iii When the asset is already obsolete due to technological advances;

iv When there are more efficient and economical solutions.

Some studies also consider the following aspects [22]:

- Availability of new technologies;
- Compliance with safety standards or other mandatory requirements;
- Availability of spare parts;
- Obsolescence that may limit the asset's use.

When the equipment enters the final phase of its LCC, it is important to determine the most rational time to withdrawal it. Several variables are important in such evaluation, including:

- Purchase price of new equipment;
- Withdrawal value;
- Operating costs;
- Maintenance costs;
- Inflation and capitalization rates.

The values of most variables are available in the asset's history, except for the withdrawal value. In this case, it is necessary to know the market value for each specific asset. Unfortunately, the knowledge of this value is difficult for many assets. In such cases, several types of depreciation can be simulated [39]:

- Linear method - devaluation is constant over the years;
- Exponential method - devaluation decreases exponentially over the years;
- Sum of digits method - devaluation is not linear over the years, but is less than exponential.

The evaluation of the equipment economic lifespan is another common method to estimate the withdrawal time: i.e., when the equipment maintenance costs exceed the cost of maintenance plus the capital amortization of new similar equipment. Farinha [22] lists three common ways to determine the economic cycle for equipment replacement:

i Income Annual Uniform Method;

ii Minimizing Total Average Cost Method;

iii Minimizing Total Average Cost Reduced to Present Value Method.

Feldens *et al.* [23] say the efficient use of physical assets is a main objective of urban passenger transport companies. In the road transport sector, the efficient use of assets is linked to a well-structured policy of fleet evaluation and replacement. Some cases of fleet replacement applied to urban buses are reported in [10, 19, 25, 30, 40, 43, 45, 51, 55]. Beichelt [10] proposes a policy for optimal scheduling replacement intervals of technical systems on the basis of a maintenance cost parameter: a system is replaced by a new one as soon as the maintenance cost within a replacement cycle reaches or exceeds a given level. Franck *et al.* [25] describe an industrial application of a Power Law Process (PLP) and a theoretical replacement model.

Pinar and Hartman [40] adopt a model for a transit fleet replacement problem with multiple types of buses. However, many cost functions are highly simplified or not based on real data, and the authors do not study the variability in vehicular characteristics, usages, and market fluctuations. The cost of replacing, refabricating, and rehabilitating buses is the focus of research by Khasnabis *et al.* [30]; they also

consider the optimal allocation proposed by the Federal Transit Administration (FTA). Other lines of research have focused on statistical analyses of fleet data and the relationships among age, utilization, and costs, [19].

Raposo *et al.* [43] present a new approach to economic models to determine the best time for bus replacement in an urban fleet. The study employs life cycle assessment and a decision support tool. Using the annual uniform income method of analysis, the authors demonstrate there is a variation in the ideal time for bus replacement.

Campos *et al.* [15] propose a generic model based on a neural stochastic process that can be applied to problems involving data with a stochastic behavior with periodic characteristics. Using neural network models, they model the behavior of an historic data series without requiring *a priori* information about the series, by generating a synthetic time series adaptable to time series. Some cases of use of neural networks and stochastic models are reported in [1, 4, 15, 24, 26, 27, 33, 35, 37, 50, 54]. Araújo and Bezerra [4] demonstrate the feasibility of a component that implements a stochastic decision support model to integrate with corporate information systems, thus contributing to the efficiency and effectiveness of the decision-making process.

Stochastic models and neural networks are often used in decision support, such as replacement [24] or interventions [35]. Other authors suggest the coordination of maintenance scheduling for the transportation fleets of many branches of a logistic service provider [27], the identification techniques of linear and nonlinear time series [33], the evaluation of vehicle fleet maintenance management indicators [50], and a chaotic time series prediction based on neural networks [54]. For work on neural networks, [15] and [26] are good references.

Other tools may contribute to the development of a new model for the optimization of bus replacement, such as Fuzzy Logic and Support Vector Machine (SVM) [14, 17, 18, 41, 44, 48, 53]. For predictive maintenance, specifically using oil analysis, several mathematical models are appropriate [2, 12, 13, 22, 46]. Lubricant degradation is not an instantaneous process; the loss of physicochemical properties and contamination are progressive over time and use of equipment, so the lubricant life is limited in service [46]. Algorithms based on exponential smoothing give interesting results [22].

3. An integrative approach to evaluate reserve fleet

The next sections present the theoretical models that support our study of the bus fleet reserve and explain their relations to maintenance policies, namely condition monitoring and predictive maintenance. A global view of the approach is the following:

- Econometric models to determine the most rational value of LCC
  - Uniform Annual Income, taking into account the following variables:
    - Operating costs
    - Maintenance costs
    - Fuel costs
  - Replacement value
  - Inflation rate
  - Capitalization rate
  - Useful life, taking into account the above variables
    - Conditioning monitoring / predictive maintenance models to maximize availability
  - Reliability KPI's

These approaches are integrated into a single analytical model that allow us to determine the best time for bus withdrawal, that is the time when Uniform Annual Income reaches its lowest value or when the equipment reaches its useful life. The size of the reserve fleet is indexed to the size of the overall fleet.

The integrated approach is valid for both new and used buses through the monetary correction effect. The analysis herein does not include technological and environmental aspects. Figure 3.1 shows the proposed integrative model to determine the size of a reserve fleet.

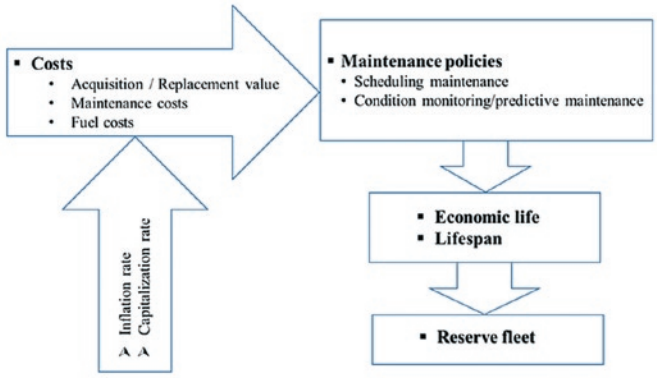


Fig. 3.1 – Integrative Model for Dimensioning a Reserve Fleet

4. Some theoretical models for asset replacement decision process

To analyze equipment replacement, two variables should be taken into account:

- Capitalization rate, i;
- Inflation rate, θ.

These rates are related in the following manner:

$$i_A = i + \theta + i\theta \tag{1}$$

where

i_A = Apparent rate.

Farinha [22] suggest several methods to determine the economic life cycle. One of these, the Annual Uniform Method (AUM), makes use of the following data:

- Cost of acquisition;
- Withdrawal value;
- Maintenance and operating costs over time;
- Apparent rate.

This paper uses the exponential method to calculate the withdrawal value of a bus; when there are no real data from the market, as is the case here, it is necessary to simulate the equipment depreciation. The exponential method seems adequate because of the high devaluation of this type of equipment over time. The formula that permit us to evaluate the annual cost depreciation can be expressed as:

$$d_l = VC_{l-1}\,(1 - \sqrt[N]{\frac{VC_N}{CA}}) \tag{2}$$

$$V_n = VC_{l-1} - d_l \tag{3}$$

where

- d_l Annual depreciation quota;
- CA Cost of Acquisition;
- N Time of life corresponding to VC_N;
- VC_N Residual value of the equipment at the end of N periods of time;
- l l=1,2,3…N;
- V_n Equipment value in period n=1,2,3…N.

The Present Net Value *per* year n (PNV_n) is expressed as follows:

$$PNV_n = CA + \sum_{j=0}^{n} \frac{CM_j + CO_j}{(1+i_A)^j} - \frac{V_n}{(1+i_A)^n} \quad (4)$$

where
- CM_j Cost of maintenance per year j= 1,2,3…n;
- CO_j Cost of operation per year j= 1,2,3…n.

The Uniform Annual Income (UAI_n) is written as follows:

$$UAI_n = \frac{i_A(1+i_A)^n}{(1+i_A)^{n-1}}(CA + \sum_{j=0}^{n} \frac{CM_j + CO_j}{(1+i_A)^j} - \frac{V_n}{(1+i_A)^n}) \quad (5)$$

$$UAI_n = \frac{i_A(1+i_A)^n}{(1+i_A)^{n-1}} NPV_n \quad (6)$$

The *UAI* indicates the time (in years) when a bus ought to be replaced. This value is equivalent to the minimum annual cost of the bus.

Another method to determine the economic cycle of equipment replacement is the Minimization of Total Average Cost Method (MTACM). This method permits the determination of the lowest average cost of bus ownership that corresponds to the most rational optimal replacement time. The capital cost and the inflation rate are not considered. The calculation procedure is as follows:

$$C'_n = \frac{1}{n}\sum_{j=1}^{n}(CM_j + CO_j) \quad (7)$$

$$C''_n = \frac{1}{n}(CA - V_n) \quad (8)$$

$$C_{n(MMTAC)} = C'_n + C''_n = \min_{n\in\{1,2,\ldots,N\}} \frac{1}{n}\left(CA - V_n + \sum_{j=1}^{n}(CM_j + CO_j)\right) \quad (9)$$

where
- n *Number of years $n \in \{1,2,3\ldots N\}$;*
- C'_n *Auxiliary variable;*
- C''_n *Auxiliary variable;*
- $C_{n(MTACM)}$ *Total average cost.*

A final option is the MTACM Reduced to Present Value (MMTAC-RPV). The calculation procedure is the same as the one above but it also considers the capital cost and inflation rate. The various maintenance and withdrawal values over time are reduced to the present value, using the following procedure:

$$C'_n = \frac{1}{n}\sum_{j=1}^{n}(\frac{CM_j + CO_j}{(1+i_A)^j}) \quad (10)$$

$$C''_n = \frac{1}{n}(CA - \frac{V_n}{(1+i_A)^n}) \quad (11)$$

$$C_{n(MMTAC-RPV)} = C'_n + C''_n = \min_{n\in\{1,2,\ldots,N\}} \frac{1}{n}\left(CA - \frac{V_n}{(1+i_A)^n} + \sum_{j=1}^{n}(\frac{CM_j + CO_j}{(1+i_A)^j})\right) \quad (12)$$

where
- $C_{n(MTACM\text{-}RPV)}$ Total Average Cost Reduced to Present Value.

An increase in business competitiveness always implies a cross dialogue between maintenance activity and economic management; therefore, it is important to consider economic variables like Return on Investment (ROI). ROI is related to greatest asset availability at the lowest cost; it determines the present value of the capital investment and cash flow corrected through the apparent rate expressed as:

$$ROI = \sum_{j=1}^{n} \frac{CF_j}{(1+i_A)^j} - CA \quad (13)$$

where
- CF_j Cash Flow;
- j j=1,2,3…n.

The Uniform Annual Income (UAI_n) and ROI can be written conjunctly as:

$$\begin{cases} UAI_n = \frac{i_A(1+i_A)^n}{(1+i_A)^{n-1}}(CA + \sum_{j=1}^{n}\frac{CM_j + CO_j}{(1+i_A)^j} - \frac{V_n}{(1+i_A)^n}) \\ ROI = \sum_{j=1}^{n}\frac{CF_j}{(1+i_A)^j} - CA \end{cases} \quad (14)$$

The Minimization of Total Average Cost Method (MMTAC) and Return on Investment (ROI) can be written conjunctly as:

$$\begin{cases} C_{n(MMTAC)} = \min_{n\in\{1,2,\ldots,N\}} \frac{1}{n}\left(CA - V_n + \sum_{j=1}^{n}(CM_j + CO_j)\right) \\ ROI = \sum_{j=1}^{n}\frac{CF_j}{(1+i_A)^j} - CA \end{cases} \quad (15)$$

The MMTAC Reduced to Present Value (MMTAC-RPV) and Return on Investment (ROI) can be written conjunctly as:

$$\begin{cases} C_{n(MMTAC-RPV)} = \min_{n\in\{1,2,\ldots,N\}} \frac{1}{n}\left(CA - \frac{V_n}{(1+i_A)^n} + \sum_{j=1}^{n}(\frac{CM_j + CO_j}{(1+i_A)^j})\right) \\ ROI = \sum_{j=1}^{n}\frac{CF_j}{(1+i_A)^j} - CA \end{cases} \quad (16)$$

To apply the models, direct operating costs and maintenance costs are taken into consideration, as are the relevant economic indicators, such as inflation and interest rates.

The study described in this paper began with a survey of the operating data for a bus fleet belonging to a medium-sized urban transport company. Based on these data, a pilot project to validate the replacement model created a simulation using a reduced number of buses. The project used historical data from 1993 to 2014. The buses were put into homogeneous groups: the buses were 21, 18, 16, 12, and 11 years old.

5. Oil analysis as a covariate for RUL prediction

There are several techniques for asset condition monitoring, like the followings: vibration analysis, thermography, visual inspection, ultrasonic measurement, and oil analysis. Among these techniques, oil analysis plays an important key role in the condition monitoring of combustion engines including the estimating of its Remaining Useful Life (RUL).

It is possible to evaluate the oil's ability to continue performing its function, and, therefore, to determine the equipment condition. The life prediction can be calculated on a statistical basis, using measurement records that permit condition forecasting so that interventions can be made before the degradation becomes severe.

Developing an effective oil analysis program requires careful planning based on a stated series of reliability goals, as is shown in Figure 5.1.

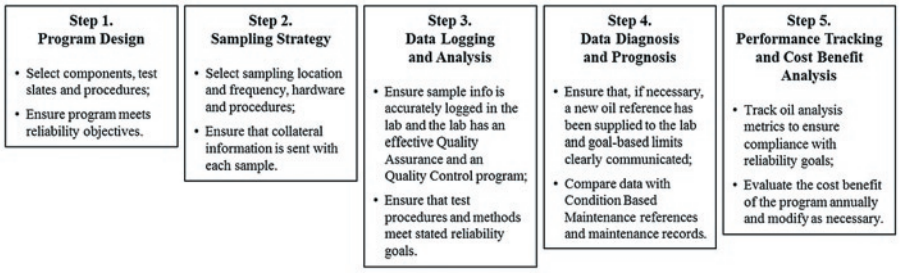


Fig. 5.1 – Reliability goals in an oil analysis program

The study described in what follows performed oil analysis on ten buses in several homogeneous groups to evaluate the importance of maintenance interventions and reserve fleet size, in three phases of analysis:

i Periodic collection of lubricant samples;

ii Study of the results, using predictive algorithms;

iii Analysis of the results using econometric replacement models.

Oil degradation was monitored in a homogeneous group of buses, with three types of oils analyzed:

i Lubricant I - 10 W 40 EHPDO (Extra High Performance Diesel Oil);

ii Lubricant II - 10 W 40 UHPDO (Ultra High Performance Diesel Oil);

iii Lubricant III - 15 W 40.

The parameters used to monitor oil degradation were the followings:

- Soot (carbonaceous material);
- Viscosity;
- Total Base Number (TBN);
- Metal wear and contamination;
- Particles.

The reference limits proposed by the laboratory data sheets were used to set the values. Important parameters turned out to be soot (%) and Iron content (ppm), as is shown below.

The model applied to monitor the degradation of soot by estimating the average value of soot was the t-student distribution. A t-student distribution was used instead of a normal distribution because of the low number of samples. The study determined whether the sample mean was different from the population mean. The estimation of the population mean, considering a tail distribution *t* and *n-1* degrees of freedom, uses the following formula:

$$\mu = \bar{X} + t_\alpha \frac{S}{\sqrt{n}} \tag{17}$$

where

- μ Population mean;
- t_α Critical t;
- $\bar{X}$ Sample mean;
- S Sample standard deviation;
- n Sample size.

The average value of the population was estimated by the significance levels of 0.001, 0.01, 0.05, 0.1 and 0.2.

The objective was to estimate the mean value of soot (threshold value = 1.5%) and iron content (threshold value = 90ppm), with the following values: 1.90% for soot and 99.80ppm for iron content. These values were evaluated above the normal limits to support conclusions of degradation in the equipment.

The analysis was complemented by an analysis of the evolution of the variables' degradation, with the Exponential Smoothing formula used to forecast their next values:

$$S_{t+1} = \alpha X_t + (1-\alpha) S_t \tag{18}$$

where

- S_{t+1} is the forecast for the next time;
- X_t is the real value recorded in the present time;
- S_t is the forecasted value for the present time;
- α is the smoothing parameter, .

The values of the variables were higher than normal limits, thus pointing to evident degradation. When these variables have high values, the equipment is at high risk and the oil must be changed.

Table 5.1 – Predicted Values for Soot (Bus X1)

Soot				
Period [km]	**Ob. Val. [%]**	**Pred. with α=0.1 [%]**	**Pred. with α=0.9 [%]**	**Pred. with α=0.5 [%]**
320 134	1.30			
330 734	1.50	1.30	1.30	1.30
355 642	1.75	1.32	1.48	1.40
376 769	1.90	1.36	1.72	1.58
472 423	2.50	1.42	1.88	1.74
500 000		1.53	2.44	2.12

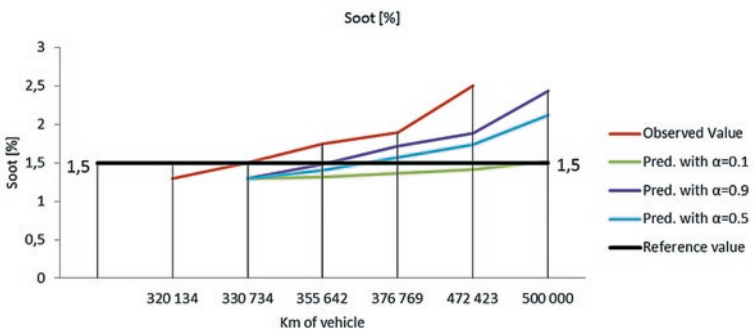


Figure 5.2 – Predicted Values for Soot (Bus X1)

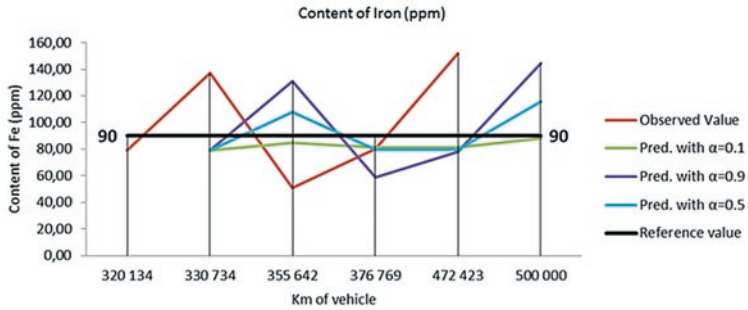


Fig. 5.3 – Predicted Values for Iron Content (Bus X1)

In Figure 5.3, the iron content increases, then suddenly decreases, dropping under the reference value. This happens when the oil is changed. The situation repeats in each utilization cycle.

Oil analysis is most effective when it is used to track metrics or benchmarks set in the planning stage. The goal is to improve the lifespan and increase MTBF by creating a performance metric that can be used to measure compliance with the stated reliability goals. Metrics provide accountability, not just for those directly involved with the oil analysis program, but for the whole plant, sending a clear message that lubrication and oil analysis are an important part of the fleet´s strategy to achieve both maintenance and performance objectives. The final stage of oil analysis is to evaluate the effectiveness of the program. This should include a cost benefit evaluation of maintenance "saves" because of oil analysis. Ongoing evaluation permits continuous improvement of the program by realigning it with either preexisting or new reliability objectives.

6. Condition based maintenance *versus* reserve fleet

The development and implementation of a policy to support strategies based on condition monitoring as described in the previous section, especially predictive maintenance, imply the consolidation of several monitoring systems. Monitoring bus "health" through condition variables will dramatically improve the optimization of the maintenance intervals, usually increasing them (the original planned preventive maintenance intervals are usually conservative), with the following implications:

- Eliminating unnecessary disassembly of parts of equipment for inspection, thereby increasing the availability of the vehicle and decreasing the maintenance costs;
- Reducing unplanned maintenance, thus increasing the availability of the vehicle and decreasing the maintenance costs;
- Reducing severe service shutdowns, increasing bus reliability and decreasing bus unavailability and immobilization costs;
- Detecting problems before they become critical, reducing intervention costs;
- Increasing the useful life of the components and the bus, thus increasing profitability.

Monitoring the condition maintenance variables may influence the reliability indicators, i.e., the MTTR and the Availability (A). The following well-known equations express the Availability, the MTTR and MTBF as a function of A:

$$A = \frac{MTBF}{MTBF + MTTR} \tag{19}$$

$$MTTR = MTBF\frac{(1-A)}{A} \tag{20}$$

$$MTBF = \frac{MTTR}{\frac{(1-A)}{A}} \tag{21}$$

Although these equations are simple, they express the relevance of the MTBF and MTTR on operational availability. In this type of vehicle (bus fleets), increasing availability by diminishing MTTR should have the direct consequence of diminishing the reserve fleet and, by consequence, the respective costs.

Figures 6.1, 6.2 and 6.3 show that when the MTTR decreases, the bus availability increases. When the MTBF increases, the bus availability also increases. Although this conclusion seems obvious, especially when the covariate oil quality is highly correlated with asset degradation, it has strategic consequences for the reserve fleet - each extra bus in the fleet represents many thousands of euros.

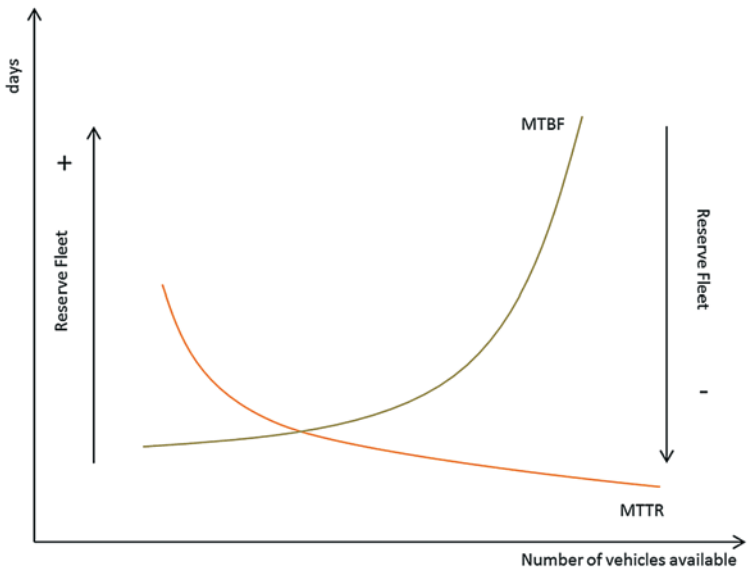


Fig. 6.1 – MTTR and MTBF versus Reserve Fleet

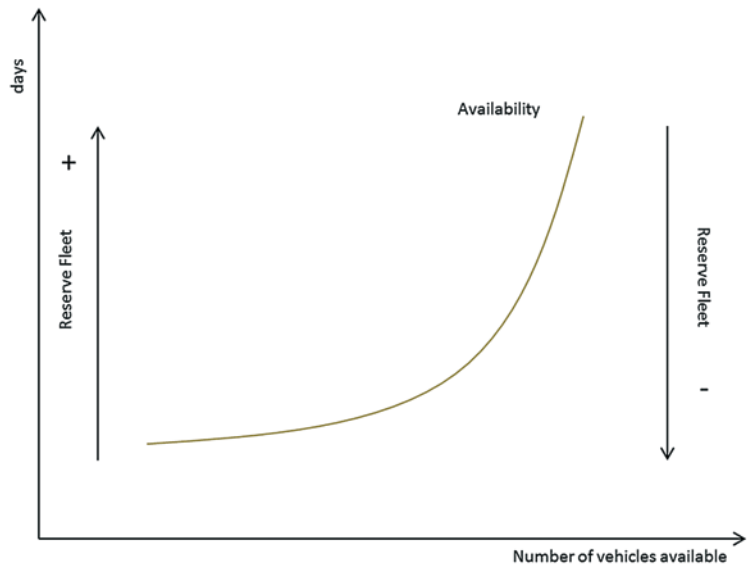


Fig. 6.2 – Availability versus Reserve Fleet

The variations in the size of the reserve fleet based on several values simulated for the MTTR are shown in Table 6.1. This table and Figure 6.3 clearly demonstrate that a policy of condition monitoring leads to a higher MTBF, a lower MTTR and, consequently, increased Availability.

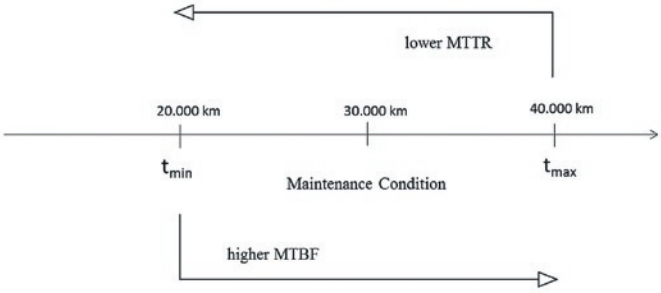


Fig. 6.3 – Maintenance Condition versus TTR/TBF (MTTR/MTBF)

Table 6.1 – MTTR versus Reserve Fleet

MTTR [days]	Bus Fleet [m]	Reserve Fleet [RF]	Interval [RF]
5	100	1,4	[1,2]
10	100	2,7	[2,3]
15	100	4,1	[4,5]
20	100	5,5	[5,6]
25	100	6,8	[6,7]
30	100	8,2	[8,9]

It is also important to demonstrate the effect of the relations of these indicators on the size of the reserve fleet. From Table 6.1, formula (22) can be inferred to determine the size of the reserve fleet:

$$RF = \frac{mMTTR}{k} \tag{22}$$

where
- *RF* Reserve fleet;
- *m* Number of fleet buses;
- *MTTR* Mean Time to Repair (days);
- *k* Number of days (in our case 365).

According to Table 6.1, the size of the reserve fleet increases with the MTTR. The lower this indicator, the lower the company's investment in a reserve fleet. The emphasis is placed on the MTTR because of its high impact on management costs, especially the indirect ones. Even a small increase in the MTTR value corresponds to an increase in the reserve fleet cost, because of the high cost of each bus, as mentioned previously.

It can be concluded that the reliability KPI's under discussion and the maintenance policies practiced by road transport companies have an enormous impact on maintenance costs and on the size of the reserve fleet. This subject is discussed in more detail in the next section.

7. Influence of MTTR on replacement time and ROI

This section discusses the influence of the MTTR on the economic cycle of bus replacement, along the bus life cycle. It also integrates the Return on Investment (ROI) of a bus with the economic cycle of bus replacement.

The Uniform Annual Income (UAI_n) Method as a function of the Mean Time to Repair (MTTR), taking into account the Return On Investment (ROI), can be written as:

$$\begin{cases} UAI_n = \dfrac{i_A(1+i_A)^n}{(1+i_A)^{n-1}}\left(CA + \sum_{j=1}^{n} \dfrac{(t\ MTTR\dfrac{CM_j}{d}) + CO_j}{(1+i_A)^j} - \dfrac{V_n}{(1+i_A)^n}\right. \\ ROI = \sum_{j=1}^{n} \dfrac{CF_j}{(1+i_A)^j} - CA \end{cases} \tag{23}$$

where
- *t* *Number of intervals considered for MTTR;*
- *d* *Number of days per year.*

The Minimization Method of Total Average Cost (MMTAC) as a function of MTTR and ROI can be written as follows:

$$\begin{cases} C_{n(MMTAC)} = \min_{n\in\{1,2,\ldots,N\}} \dfrac{1}{n}\left(CA - V_n + \sum_{j=1}^{n}\left((tMTTR\dfrac{CM_j}{d}) + CO_j\right)\right) \\ ROI = \sum_{j=1}^{n} \dfrac{CF_j}{(1+i_A)^j} - CA \end{cases} \tag{24}$$

The MMTAC Reduced to Present Value (MMTAC-RPV) in function of MTTR and ROI can be written as:

$$\begin{cases} C_{n(MMTAC-RPV)} = \min_{n\in\{1,2,\ldots,N\}} \dfrac{1}{n}\left(CA - \dfrac{V_n}{(1+i_A)^n} + \sum_{j=1}^{n}\left(\dfrac{(tMTTR\dfrac{CM_j}{d}) + CO_j}{(1+i_A)^j}\right)\right) \\ ROI = \sum_{j=1}^{n} \dfrac{CF_j}{(1+i_A)^j} - CA \end{cases} \tag{25}$$

The paper has emphasized the importance of condition monitoring, its relations with the MTTR and the importance of this KPI to the time of bus replacement. An example of the influence of this ratio, using the Uniform Annual Income method, is given in Table 7.1 and Figure 7.1. The table and the figure clearly show the effect of the

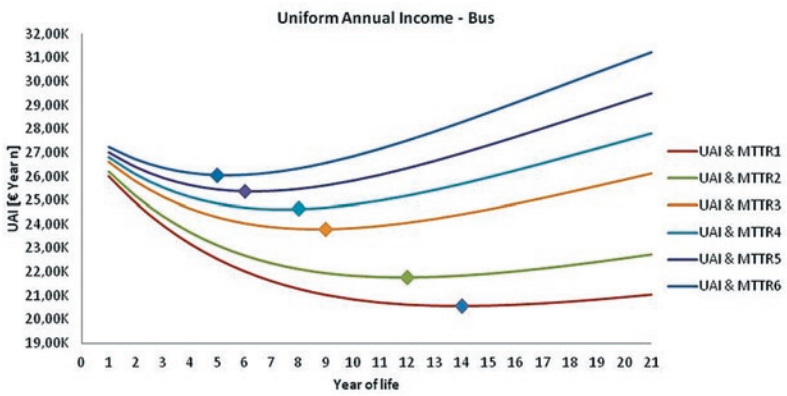


Figure 7.1 – Influence of MTTR on UAI - ROI

Table 7.1 – Influence of MTTR on UAI

Vehicles										UAI [€ Year n]					
Year	Year j	CA [€]	i_A [%]	$MTTR_1$	$MTTR_2$	$MTTR_3$	$MTTR_4$	$MTTR_5$	$MTTR_6$	UAI & $MTTR_1$	UAI & $MTTR_2$	UAI & $MTTR_3$	UAI & $MTTR_4$	UAI & $MTTR_5$	UAI & $MTTR_6$
1993	0	110.66K	8%												
1994	1		8%	5	10	20	25	30	35	25.99K	26.20K	26.61K	26.81K	27.02K	27.22K
1995	2		8%	5	10	20	25	30	35	24.88K	25.18K	25.79K	26.09K	26.40K	26.70K
1996	3		8%	5	10	20	25	30	35	23.94K	24.34K	25.14K	25.54K	25.95K	26.35K
1997	4		8%	5	10	20	25	30	35	23.17K	23.67K	24.65K	25.15K	25.64K	26.13K
1998	5		8%	5	10	20	25	30	35	22.54K	23.12K	24.29K	24.87K	25.46K	26.04K
1999	6		8%	5	10	20	25	30	35	22.02K	22.69K	24.04K	21.71K	25.38K	26.05K
2000	7		8%	5	10	20	25	30	35	21.61K	22.36K	23.88K	24.64K	25.39K	26.15K
2001	8		8%	5	10	20	25	30	35	21.28K	22.12K	23.80K	24.64K	25.48K	26.32K
2002	9		8%	5	10	20	25	30	35	21.03K	21.95K	23.79K	24.71K	25.63K	26.55K
2003	10		8%	5	10	20	25	30	35	20.84K	21.84K	23.83K	24.83K	25.83K	26.83K
2004	11		8%	5	10	20	25	30	35	20.70K	21.78K	23,92K	25.00K	26.07K	27.15K
2005	12		8%	5	10	20	25	30	35	20.62K	21.76K	24.06K	25.20K	26.35K	27.49K
2006	13		8%	5	10	20	25	30	35	20.57K	21.79K	24.22K	25.44K	26.65K	27.87K
2007	14		8%	5	10	20	25	30	35	20.56K	21.84K	24.41K	25.69K	26.98K	28.26K
2008	15		8%	5	10	20	25	30	35	20.57K	21.92K	24.62K	25.97K	27.32K	28.67K
2009	16		8%	5	10	20	25	30	35	20.61K	22.02K	24.85K	26.26K	27.67K	29.09K
2010	17		8%	5	10	20	25	30	35	20.67K	22.14K	25.09K	26.56K	28.04K	29.51K
2011	18		8%	5	10	20	25	30	35	20.74K	22.28K	25.34K	26.87K	28.40K	29.94K
2012	19		8%	5	10	20	25	30	35	20.83K	22.42K	25.60K	27.18K	28.77K	30.36K
2013	20		8%	5	10	20	25	30	35	20.93K	22.57K	25.86K	27.50K	29.14K	30.79K
2014	21		8%	5	10	20	25	30	35	21.04K	22.73K	26.12K	27.82K	29.51K	31.21K

Table 7.2 – MTTR versus UAI – ROI

Vehicles										UAI [€ Year n]						ROI
Year	Year j	CA [€]	i_A [%]	$MTTR_1$	$MTTR_2$	$MTTR_3$	$MTTR_4$	$MTTR_5$	$MTTR_6$	UAI & $MTTR_1$	UAI & $MTTR_2$	UAI & $MTTR_3$	UAI & $MTTR_4$	UAI & $MTTR_5$	UAI & $MTTR_6$	
1993	0	110.66K	8%													
1994	1		8%	5	10	20	25	30	35	25.99K	26.20K	26.61K	26.81K	27.02K	27.22K	-87.51K
1995	2		8%	5	10	20	25	30	35	24.88K	25.18K	25.79K	26.09K	26.40K	26.70K	-70.36K
1996	3		8%	5	10	20	25	30	35	23.94K	24.34K	25.14K	25.54K	25.95K	26.35K	-54.57K
1997	4		8%	5	10	20	25	30	35	23.17K	23.67K	24.65K	25.15K	25.64K	26.13K	-40.01K
1998	5		8%	5	10	20	25	30	35	22.54K	23.12K	24.29K	24.87K	25.46K	26.04K	-26.61K
1999	6		8%	5	10	20	25	30	35	22.02K	22.69K	24.04K	21.71K	25.38K	26.05K	-14.25K
2000	7		8%	5	10	20	25	30	35	21.61K	22.36K	23.88K	24.64K	25.39K	26.15K	-2.88K
2001	8		8%	5	10	20	25	30	35	21.28K	22.12K	23.80K	24.64K	25.48K	26.32K	7.61K
2002	9		8%	5	10	20	25	30	35	21.03K	21.95K	23.79K	24.71K	25.63K	26.55K	17.26K
2003	10		8%	5	10	20	25	30	35	20.84K	21.84K	23.83K	24.83K	25.83K	26.83K	26.15K
2004	11		8%	5	10	20	25	30	35	20.70K	21.78K	23,92K	25.00K	26.07K	27.15K	34.35K
2005	12		8%	5	10	20	25	30	35	20.62K	21.76K	24.06K	25.20K	26.35K	27.49K	41.89K
2006	13		8%	5	10	20	25	30	35	20.57K	21.79K	24.22K	25.44K	26.65K	27.87K	48.84K
2007	14		8%	5	10	20	25	30	35	20.56K	21.84K	24.41K	25.69K	26.98K	28.26K	55.24K
2008	15		8%	5	10	20	25	30	35	20.57K	21.92K	24.62K	25.97K	27.32K	28.67K	61.14K
2009	16		8%	5	10	20	25	30	35	20.61K	22.02K	24.85K	26.26K	27.67K	29.09K	66.56K
2010	17		8%	5	10	20	25	30	35	20.67K	22.14K	25.09K	26.56K	28.04K	29.51K	71.56K
2011	18		8%	5	10	20	25	30	35	20.74K	22.28K	25.34K	26.87K	28.40K	29.94K	76.17K
2012	19		8%	5	10	20	25	30	35	20.83K	22.42K	25.60K	27.18K	28.77K	30.36K	80.41K
2013	20		8%	5	10	20	25	30	35	20.93K	22.57K	25.86K	27.50K	29.14K	30.79K	84.31K
2014	21		8%	5	10	20	25	30	35	21.04K	22.73K	26.12K	27.82K	29.51K	31.21K	87.91K

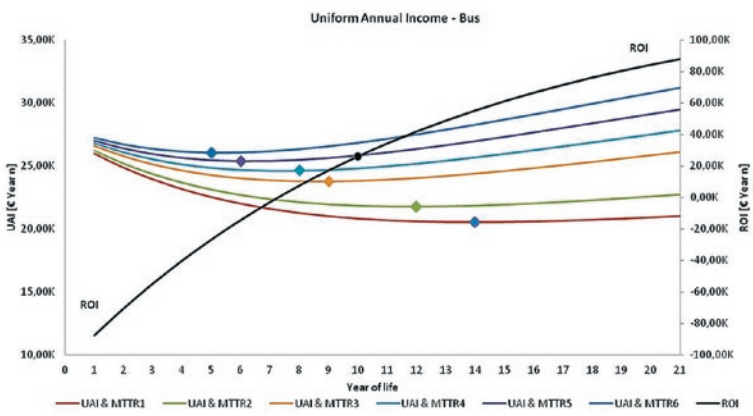


Figure 7.2 – Influence of MTTR on UAI – ROI

MTTR on the bus replacement time. Note: the MTTR values of 5, 10, 20, 25, 30 and 35 days were used for the simulation.

Table 7.1 and Figure 7.1 verify the influence of the MTTR in the econometric models. An increase or decrease in the MTTR over time causes a variation on the withdrawal point. The replacement point can vary by more than nine years: i.e. the time for replacement is five years, but if the MTTR decreases, the replacement time becomes 14 years. This shows the importance of this KPI in the management of a bus fleet; it affects the size of the reserve fleet and also the cost.

In addition, it is useful to analyse the relevance of economic ratios to the evaluation of investments, such as the ROI. Table 7.2 and Figure 7.2 show the influence of the MTTR on the calculation of Uniform Annual Income (UAI_n) and on the ROI of a bus. If the UAI curves are compared with the ROI, it becomes obvious that the smaller the UAI, the greater the company's profit.

Table 7.3 and Figure 7.3 focus on a MTTR of five days. It can be noted that for MTTR = 5 days, the replacement time is 14 years, and the value of the UAI is 20.56K€. Note that in the 10th year, the ROI value is 26.15K€. Therefore, the ROI is higher than the value of UAI_n, and the equipment has started yielding profit to the company. The period between the year the equipment starts giving a profit (year=10) and the year it is replaced corresponds to the year when the UAI is minimal (year=14), with a profit value of 103.19K€.

Table 7.3 – Influence of MTTR=5 on UAI with ROI

Vehicles					UAI [€ Year n]	ROI
Year	Year j	CA [€]	i_A [%]	$MTTR_1$	UAI & $MTTR_1$	
1993	0	110.66K	8%			
1994	1		8%	5	25.99K	-87.51K
1995	2		8%	5	24.88K	-70.36K
1996	3		8%	5	23.94K	-54.57K
1997	4		8%	5	23.17K	-40.01K
1998	5		8%	5	22.54K	-26.61K
1999	6		8%	5	22.02K	-14.25K
2000	7		8%	5	21.61K	-2.88K
2001	8		8%	5	21.28K	7.61K
2002	9		8%	5	21.03K	17.26K
2003	10		8%	5	20.84K	26.15K
2004	11		8%	5	20.70K	34.35K
2005	12		8%	5	20.62K	41.89K
2006	13		8%	5	20.57K	48.84K
2007	14		8%	5	20.56K	55.24K
2008	15		8%	5	20.57K	61.14K
2009	16		8%	5	20.61K	66.56K
2010	17		8%	5	20.67K	71.56K
2011	18		8%	5	20.74K	76.17K
2012	19		8%	5	20.83K	80.41K
2013	20		8%	5	20.93K	84.31K
2014	21		8%	5	21.04K	87.91K

Table 7.4 – Influence of MTTR=5 on UAI with ROI

Vehicles					UAI [€ Year n]	ROI
Year	Year j	CA [€]	i_A [%]	$MTTR_1$	UAI & $MTTR_1$	
1993	0	110.66K	8%			
1994	1		8%	5	25.99K	-64.36K
1995	2		8%	5	24.88K	-21.41K
1996	3		8%	5	23.94K	18.44K
1997	4		8%	5	23.17K	55.41K
1998	5		8%	5	22.54K	89.71K
1999	6		8%	5	22.02K	121.54K
2000	7		8%	5	21.61K	151.06K
2001	8		8%	5	21.28K	178.45K
2002	9		8%	5	21.03K	203.80K
2003	10		8%	5	20.84K	227.44K
2004	11		8%	5	20.70K	249.32K
2005	12		8%	5	20.62K	269.61K
2006	13		8%	5	20.57K	288.44K
2007	14		8%	5	20.56K	305.90K
2008	15		8%	5	20.57K	322.10K
2009	16		8%	5	20.61K	337.14K
2010	17		8%	5	20.67K	348.38K
2011	18		8%	5	20.74K	358.82K
2012	19		8%	5	20.83K	368.50K
2013	20		8%	5	20.93K	377.49K
2014	21		8%	5	21.04K	385.83K

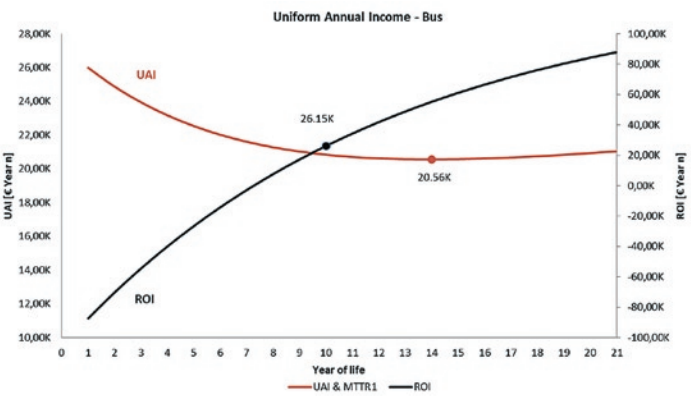


Figure 7.3 – Influence of MTTR=5 on UAI with ROI

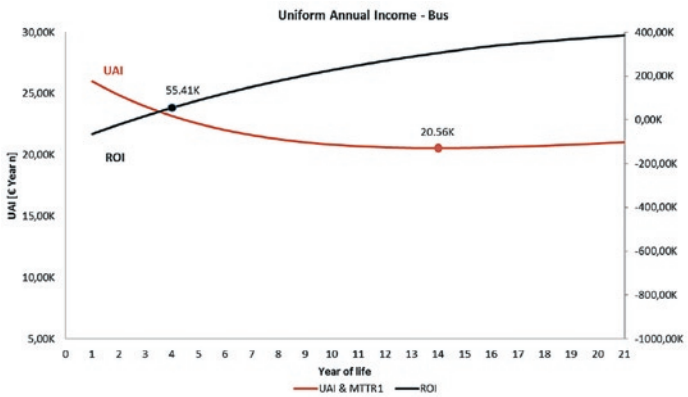


Figure 7.4 – Influence of MTTR=5 on UAI with ROI

The example shown in Table 7.4 and Figure 7.4 uses data from the previous example, but with a higher value of ROI per year. The higher ROI is used to show its relevance to company profit when other parameter values remain the same.

The ROI value is higher each year, and the bus starts to give a profit in four years with an ROI value of 55.41K€. Then, for the period from the time the equipment starts to give profit (year=4) to the time it is replaced - when the Uniform Annual Income is minimal (=14 years) - the total profit generated by the bus is 1905.76K€.

8. Conclusions

The paper gives an overview of ongoing research in the use of econometric models to determine the optimal time to replace a bus. Indicators associated with acquisition, maintenance and operation costs, among others, guide the economic aspects. The paper demonstrates that the variations in the most rational time to replace a bus are influenced by the MTTR and the ROI.

Another aspect emphasized in the paper is the importance of implementing condition based maintenance policies; this helps to reduce the costs associated with maintenance and increases availability. To make the point, the paper uses the example of oil analysis.

The paper proposes integrating several methodologies in an econometric global model that supports the decision to replace buses and to calculate the size of the reserve fleet. It validates these methodologies using real data from a bus fleet. The results show the value of the integrated model.

It describes simulations using real data that verify the relations among the maintenance policy, the time of replacement and the dimension of the reserve fleet.

An integrated econometric global model to determine the time of bus replacement and the size of the reserve fleet based on a condition monitoring maintenance model can aid management decision making in these areas.

References

1. Amaya E J, Tonaco R, Souza R Q, Álvares A J. Intelligent System Maintenance Based on Condition for Balbina Dam. University of Brasilia, Department of Mechanical Engineering and Mechatronics, Innovation Group in Industrial Automation (GIAI), CEP 70910-900, 8th Congress Engineering Mechanics Ibero Americano, Brasília, DF, Brazil, 2007.
2. André J C S. Probability and Statistics for Engineering. Lisbon: 1st Editions, Lidel - Technical Issues Lda, 2008; ISBN 9789727574773.
3. Aoudia M, Belmokhtar O. Economic impact of maintenance management ineffectiveness of an oil gas company. Journal of Quality in Maintenance Engineering 2008; 14(3): 237-261, https://doi.org/10.1108/13552510810899454.
4. Araujo M S, Bezerra C A. Development of components for stochastic systems for decision support. PUCPR Brazilian, Congress of Computer Science, Software Engineering, Brazil, 2004; 101-107.
5. Assaf Neto A. Corporate finance and value. São Paulo, Atlas, 2005; ISBN: 9788522460144.
6. Assis R, Julião J. Maintenance Management or Asset Management? (Costs over the life cycle). Communication 10th National Congress Maintenance, APMI, Figueira da Foz, Portugal, 2009.
7. Assis R. Decision support in maintenance management of physical assets. Lisbon: 1st Edition, Lidel - Technical issues, Lda, 2010. ISBN: 9789897521126.
8. ASTM International - Standard practice for measuring life-cycle costs of buildings and building system. Annual Book of ASTM Standards: ASTM International West Conshohocken, PA, E 917, 2002; 4(11).
9. BAS PAS 55 - Asset Management: PAS 55-1, Part 1: Specification for the optimized management of physical assets | PAS 55-2, Part 2: Guidelines for the application of PAS 55-1. British Standards, UK, 2008.
10. Beichelt F. A replacement policy based on limiting the cumulative maintenance cost. Department of Statistics and Acturial Science, University of Witwatersrand, Johannesburg, South Africa; International Journal of Quality & Reliability Management. MCB University Press, 0265-671X, 2001; 18(1): 76-83.
11. Bescherer F. Established Life Cycle Concepts in the Business Environment – Introduction and terminology, Laboratory of Industrial Management Report Series, report 1/2005, Helsinki University, 2005; ISBN 951-22-7549.
12. Cabral J S. Organization and Maintenance Management. Lisbon: 6th Edition, Lidel - Technical Issues Lda, 2006; ISBN: 9789727574407.
13. Cabrita C P, Cardoso A J M. Concepts and definitions of failure and breakdown in the Portuguese maintenance standards NP EN 13306: 2007 and NP EN 15341: 2009. CISE - Electromechatronic Systems Research Centre, University of Beira Interior, 17 Ibero-American Congress on Maintenance, Cascais, Portugal, 2013.
14. Campello R J G B, Amaral W C. Modelling And Linguistic Knowledge Extration From Systems Using Fuzzy Relation Models, Fuzzy Sets and Systems 2001; 121: 113-126, https://doi.org/10.1016/S0165-0114(99)00175-X.
15, Campos L C D, Vellasco M M B R, Lazo J G L. A stochastic model based on neural networks. PhD thesis, UFJF, Juiz de Fora, Brazil, 2010.
16. Casarotto F N. Investment analysis - financial mathematics, economic engineering, decision making, business strategy. São Paulo, Atlas, 2000; ISBN: 85-224-2572-8.
17. Chena D, Wanga L, Li L. Position computation models for high-speed train based on support vector machine approach. Control and Safety, Beijing Jiaotong University, Beijing 100044, China, 2015, https://doi.org/10.1016/j.asoc.2015.01.017.
18. Couellan N, Jana S, Jorquera T, George J P. Self-adaptive Support Vector Machine: A multi-agent optimisation perspective. Université de Toulouse, UPS IMT, F-31062 Toulouse Cedex 9, France, 2015, https://doi.org/10.1016/j.eswa.2015.01.028.
19. Di J, Hauke L. Optimal fleet utilization and replacement. Transportation Research Part E, 2000; 36(1): 3–30. ISSN: 1366-5545.
20. Durairaj S K, et al. Evaluation of life cycle cost analysis methodologies, corporate Environment Strategy 2002; 9(1): 30-39.
21. Emblemsvag J. Activity-based life-cycle costing. Managerial Auditing Journal 2001; 16(1): 17-27, https://doi.org/10.1108/02686900110363447.
22. Farinha J M T. Maintenance - Terminology and New Management Tools. Lisbon: 1st Edition, Monitor - Design and Publishing, Lda, 2011; ISBN 978-972-9413-82-7.
23. Feldens A G, Muller C J, Filomena T P, Neto F J K, Castro A S, Anzanello M J. Policy Assessment and Fleet Replacement by Means of Model Multicriteria Adoption. Porto Alegre, Brazil, 2010; ISSN 1980-4814.
24. Figueiredo L M J. Multicriteria model to support the replacement of hospital medical equipment, PhD thesis, IST, Lisbon, Portugal, 2009.
25. Francis K N, Leung and Ada, Cheng L M. Determining replacement policies for bus engines. City University of Hong Kong, Hong Kong; International Journal of Quality & Reliability Management, MCB University Press, 0265-671X, 2000; 17 (7): 771-783.
26. Gurney K. An introduction to neural networks, London, UCL Press, 1997; ISBN: 1857285034, https://doi.org/10.4324/9780203451519.
27. Huang J, Yao M. On the coordination of maintenance scheduling for transportation fleets of many branches of a logistic service provider. Ling Tung University1 Ling Tung Road, Nantun, Taichung 408, Taiwan, ROC, 2008, https://doi.org/10.1016/j.camwa.2008.01.037.
28. ISO 55000:2014 - Asset management - Overview, principles and terminology; ISO 55001:2014 - Asset management - Management systems – Requirements; ISO 55002:2014 - Asset management - Management systems - Guidelines for the application of ISO 55001, 2014.
29. Jennifer L R, Joseph C H. Equipment replacement under continuous and discontinuous technological change. IMA Journal of Management Mathematics 2005; 16 (1): 23-36, https://doi.org/10.1093/imaman/dph027.
30. Khasnabis S, Alsaidi E, Ellis R. Optimal allocation of resources to meet transit fleet requirements. Journal of Transportation Engineering 2002; 128 (6): 509-518, https://doi.org/10.1061/(ASCE)0733-947X(2002)128:6(509).
31. Korpi E, Ala-Risku T. Life cycle costing: a review of public case studies. Managerial Auditing Journal 2008; 23 (3): 240-261, https://doi.org/10.1108/02686900810857703.
32. Lindholm A, Suomala P. The possibilities of Life Cycle Costing in Outsourcing Decision Making. Frontiers of E-Business Research 2004; 226–241.
33. Luna I, Ballini R, Soares S. Identification technique of linear and nonlinear time series. Magazine Automation and Control 2006; 17 (3): 245-256.
34. Makridakis S, Wheelwright S, Hyndman R F. Forecasting -Methods and Applications. New York: John Wiley & Sons, 1998; ISBN 0-471-53233-9.
35. Marco A R, Angelo A D, Leizer S, Silvio A B V. The use of Bayesian networks in the decision making process interventions in equipment. Industrial Engineering Program, Federal University of Bahia, Polytechnic School, XVIII Brazilian Congress Auto / 12 to 16 September 2010

Federation, 40210-630, Salvador, Brazil, 2010; 5058-5064.
36. Motta R R, Calôba G M. Investment analysis: decision making in industrial projects. São Paulo: Atlas, 2002; ISBN: 9788522430796.
37. Múller D. Stochastic processes and applications, Collection Economic volume - 2nd Series off Financial Law and Tax. Almedina, 2007.
38. Natali H, Yuri Y. Optimal equipment replacement without paradoxes: A continuous analysis. Operations Research Letters 2007; 35 (2): 245–250, https://doi.org/10.1016/j.orl.2006.03.001.
39. Oliveira J A N. Economic Engineering - an approach to investment decisions. São Paulo: McGraw-Hill Brazil, 1982.
40. Pinar K, Hartman J. Case Study: Bus Fleet Replacement. The Engineering Economist 2004; 49 (3): 253-278, https://doi.org/10.1080/00137910490498951.
41. Pooyan N, Shahbazian M, Salahshoor K, Hadian M. Simultaneous Fault Diagnosis using multi class support vector machine in a Dew Point process. Department of Instrumentation and Automation, Petroleum University of Technology, Ahwaz, Iran, 2015, https://doi.org/10.1016/j.jngse.2015.01.043.
42. Portuguese NP EN 15341: 2009 - Maintenance - performance indicators of maintenance (KPI), 2009.
43. Raposo H D N, Farinha J T, Oliveira R, Ferreira L A, André J. Time Replacement Optimization Models for Urban Transportation Buses with Indexation to Fleet Reserve. MPMM – Maintenance Performance Measurement and Management; Coimbra, Portugal, 2014; 1(1): ISBN 978-972-8954-43-7, https://doi.org/10.14195/978-972-8954-42-2_7.
44. Romaniuk M. On simulation of maintenance costs for water distribution system with fuzzy parameters. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (4): 514–527, https://doi.org/10.17531/ein.2016.4.6.
45. Scarf P A, Bouamra O A. Capital equipment replacement model for a fleet with variable size. Centre for OR and Applied Statistics, University of Salford, UK, Journal of Quality in Maintenance Engineering, © MCB University Press, 1355-251, 1999; 5 (1): 40-49, https://doi.org/10.1108/13552519910257050.
46. Seabra J, Graça B. Analysis of oils and greases in service. Proceedings of the Fifth National Congress of Industrial Maintenance - APMI, Figueira da Foz, 1996.
47. Simões A S. Conditional maintenance to Pollutant Emissions in Urban Buses, Using Degradation Prediction Models Based on the technology of vehicles and the Operating Conditions; PhD final thesis. Instituto Superior Técnico, 2011.
48. Tsoukalas L H, Uhrig R E. Fuzzy and neural approaches in engineering, New York: John Wiley, 1996. ISBN: 0471160032.
49. Vey I H, Rosa R M. Fleet replacement in municipal passenger transportation company: a case study. Federal University of Santa Maria, Electronic Accounting Journal 2004; 1 (1): 150–173.
50. Vujanovic D, Momcˇilovic V, Bojovic N, Papic V. Evaluation of vehicle fleet maintenance management indicators by application of DEMATEL and ANP. University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, Belgrade, Serbia, 2012, https://doi.org/10.1016/j.eswa.2012.02.159.
51. Wijaya A R, Lundberg J, Kumar U. Robust-optimum multi-attribute age-based replacement policy. Journal of Quality in Maintenance Engineering 2012; 18 (3): 325-343.
52. William G Sullivan, Thomas N McDonald, Eileen M Van Aken. Robotics and Computer Integrated Manufacturing 2002; 18 (3–4): 255–265, https://doi.org/10.1016/S0736-5845(02)00016-9.
53. Yager R R, Zadeh L A. Introduction to fuzzy logic applications-An in intelligent systems Boston, Kluwer Academic, Publishers, 1992.
54. Zhao H. A chaotic time series prediction based on neural network: Evidence from the Shanghai composite index in China. Test and Measurement 2009; 2 (1): 382 –385.
55. Zohrul, Kabir A B M. Evaluation of overhaul/replacement policy for a fleet of buses. Journal of Quality in Maintenance Engineering 1996; 2 (3): 49-59 ,https://doi.org/10.1108/13552519610130440.

Hugo RAPOSO
CEMMPRE - Centre for Mechanical Engineering, Materials and Processes.
University of Coimbra, 3030-788 Coimbra, Portugal

José Torres FARINHA
CEMMPRE - Centre for Mechanical Engineering, Materials and Processes.
University of Coimbra, 3030-788 Coimbra, Portugal
and
IPC - Polytechnic Institute of Coimbra.
3000-271 Coimbra, Portugal

Luís FERREIRA
UP – University of Porto
4200-465 Porto, Portugal

Diego GALAR
Lulea University of Technology, Sweden
Division of Operation and Maintenance Engineering, Department of Civil, Environmental and Natural Resources Engineering, 971- 87 Luleå, Sweden

E-mails: hugrap@gmail.com, torres.farinha@dem.uc.pt, lferreir@fe.up.pt, diego.galar@ltu.se

Article citation info:
SURBLYS V, ŽURAULIS V, SOKOLOVSKIJ E. Estimation of road roughness from data of on-vehicle mounted sensors. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 369–374, http://dx.doi.org/10.17531/ein.2017.3.7.

Vytenis SURBLYS
Vidas ŽURAULIS
Edgar SOKOLOVSKIJ

ESTIMATION OF ROAD ROUGHNESS FROM DATA OF ON-VEHICLE MOUNTED SENSORS

OCENA NIERÓWNOŚCI DROGOWYCH ZGODNIE Z DANYMI CZUJNIKÓW ZAINSTALOWANYCH NA POJEŹDZIE

In this work, methods of estimation of road roughness by processing data of laser sensors of distance measurement are investigated. Depending on the layout of the sensors and the processing of the data, the Japanese, symmetrical and asymmetrical roughness estimation methods are elaborated in this work. The experimental research that was carried out by "Japanese method" on two known profile road surfaces allowed evaluating the influence of the vehicle's suspension and the body pitch rate oscillation on the calculation of the roughness. The validation of the results showed that the compensation of the suspension travel does not warrant the sufficient accuracy of the estimation of the road roughness because the pivoting of the bodywork has a greater influence on the processing of the signals recorded by the sensors attached to it.

***Keywords**: road roughness, laser sensor, pitch angle, suspension deflection.*

Artykuł przedstawia i ocenia metody określania nierówności nawierzchni drogi na podstawie danych z laserowych czujników pomiaru odległości. W artykule omówiono opracowane w Japonii symetryczne i asymetryczne metody określania nierówności nawierzchni, które różnią się sposobem rozmieszczenia czujników i przetwarzania danych. Badania eksperymentalne przeprowadzone na podstawie „metody japońskiej" na dwóch zadanych profilach nawierzchni pozwoliły na ocenę wpływu zawieszenia samochodu i zmian przechyłu nadwozia względem osi podłużnej na obliczanie nierówności nawierzchni. Walidacja wyników pokazała, że kompensacja odchylenia zawieszenia nie zapewnia wystarczającej dokładności pomiaru nierówności drogi, ponieważ przechyły nadwozia mają większy wpływ na przetwarzanie sygnałów rejestrowanych przez zamontowane na nim czujniki.

***Słowa kluczowe**: nierówności drogi, czujnik laserowy, przechył (kąt) względem osi podłużnej, odchylenie zawieszenia.*

1. Introduction

The road surface profile is one of the main parameters which influences the vehicle's road holding and comfort characteristics. When the car moves on a rough road surface, the recorded parameters of the road surface profile can be applied in improving the dynamic characteristics. One of the methods of improving the road holding is the changing of the suspension damping force by active or semi-active type shock absorbers. That would ensure better driving comfort, vehicle steerability and the overall road safety [11, 25]. Then using active or semi-active suspension road irregularities are recorded then front wheel comes on bump or pothole. Therefore, road irregularities measurement with laser sensors lets preview roughness before wheel comes on it. It lets innovatively control suspension and gets better road holding and comfort characteristics.

Many methods are designed in order to determine the road disturbances. Some methods are based by the profilograph measuring device, others by profilometer [9]. Both methods are very expensive and are used only in road building and maintenance companies [5]. Also, a profilometer which consists of 30 aluminium arms is designed. The device measures a 3 m wide road surface and the measurements are done every 10.18 mm with the help of the arms [3]. It is widely used dynamometer trailer with measuring single wheel named SRT-4. Equipment is manufactured at Road and Bridge Research Institute, Poland [16]. With this trailer is able to investigate road surface adhesion and automotive tyre anti-slip properties [17].

A number of studies carried out to estimate road profile with factory-mounted sensors and using same filters and mathematical algorithms. Widely used Youla–Kučera parameterization technique also known as Q-parameterization [6]; Kalman filter [18]; Extended Kalman filter [4] and etc. Mathematical algorithms require very fast processing, because it requires additional calculations and precision still unmatched by direct measurement (laser, stereo camera, etc.).

Furthermore, there are methods to determine the road disturbances in standard cars by using additional devices, e.g., stereo camera which can scan the road surface and also forecast the disturbances by using the available sensors that are mounted in the car [5, 7].

Another method used to evaluate the road roughness is the International Roughness Index (IRI) [13, 14, 20]. IRI is an index defined by the number of disturbances, their size and layout per unit length of the road surface [22]. Scientist P. Mučka [14] is summarised IRI limit values for new, reconstructed, or rehabilitated roads and road classification schemes used around the world.

IRI index is calculated by simulating the quarter-car model, also known as *Golden car*, moving on the measured road profile at a speed of 80 km/h. The model allows to measure the displacements of the suspension and wheel, compensate them and evaluate the road profile. Also index could be measured using an instrumented vehicle equipped with non-contact sensors (infrared, laser or acoustic etc.) and accelerometers [1]. Different IRI values are supposed for different surfaces (Fig. 1). The measurement units for IRI are m/km, in/mi.

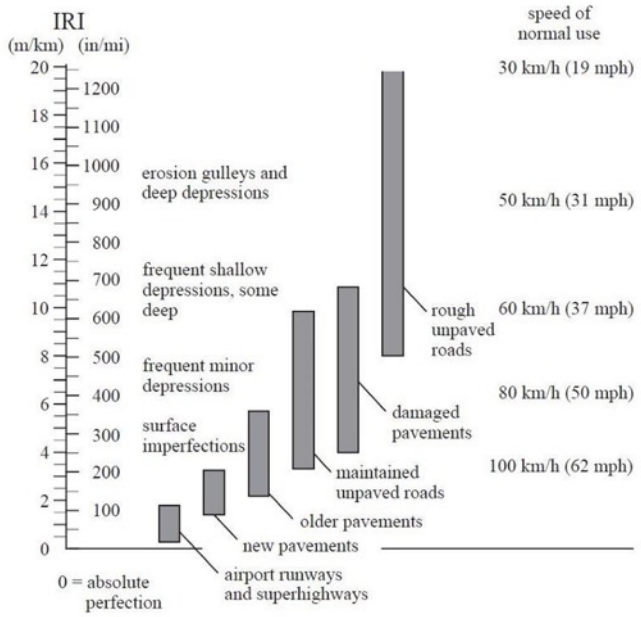


Fig. 1. IRI meanings depending on the road surface [20]

Another index describing the road disturbances is half-car roughness index (HRI). When determining this index, two wheel tracks and a dynamic model of the "half-car" are used. The IRI index correlates with the HRI index at an approximate ratio $HRI \approx 0.8\ IRI$ [19].

The aim of this work – after reviewing the road roughness estimation methods, to experimentally determine the road profile using the Japanese method and by applying the suspensions deflection and the body longitudinal roll compensation to improve the Japanese method.

2. Review of existing methods

Yuan et. al. [24] describes a method how to determine the road surface disturbances using laser line detection devices. The influence of the car and relief shadow on the accuracy of the measurement is described. It is stated that for the use of this method the road surface has to be clear. In the article [2], this method is applied to platooning of trucks. The first truck measures the road surface by lasers and transfers the obtained parameters to the following trucks. For the processing of the results, the *Naïve Bayes* method is used [12].

Another method to determine the road disturbances is to use the data of the car's acceleration [8, 15]. During the moving of the car, the data from accelerometer are recorded, they are optimised and, based on The Cross Entropy theory, the road disturbances are forecasted. Many cars have factory mounted accelerometers because the active safety systems of the car use deceleration and acceleration meanings. The research which was carried out with 5 road surfaces showed that the forecasting of the road surface from the acceleration parameters is sufficiently accurate (error up to ±2%).

2.1. Japanese method

In Japan, in order to determine the road disturbances, a method measuring the height differences from the road surface is used (Fig. 2). The measurements are done every 1.5 m at three points. While doing the measurement, all the results are recorded, and to get better accuracy, lasers are used instead of rolls [21].

The recorded measurement results are processed using equation:

$$d_i = h_i - \frac{1}{2}(h_{i-1} + h_{i+1}) \tag{1}$$

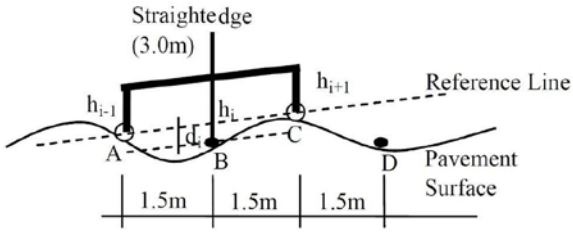


Fig. 2. The road roughness measurement method used in Japan [21]

Here d_i – the roughness of the road profile; h_i, h_{i-1}, h_{i+1} – the distances measured at characteristic points.

From the Fig. 2 It can be seen that after determining the roughness of the road profile d_i at the point B, the measurement point is shifted to position D, where the heights h_i, h_{i-1}, h_{i+1} are respectively at the points B, C ir D. The baseline becomes a line, going through the points B, C and D. Using the equation (1) d_i can be calculated at the point C. After completing a group of measurements, positive and negative bump values are obtained. These bumps are height differences based on the baseline every 1.5 m.

The standard deflection of the longitudinal roughness is calculated by summing the road profile heights by using the equation:

$$\sigma = \sqrt{\frac{n_r \sum d_i^2 - (\sum d_i)^2}{n_r(n_r - 1)}} \tag{2}$$

Here σ – the standard deflection of the longitudinal roughness (mm); d_i– the roughness of the road profile; n_r – the number of the recorded date.

Japan Highway Public Corporation recommends to calculate the coefficient σ from the recorded values of a road section of length of 150 m.

2.2. Symmetrical estimation method

Two more methods for estimation of the road disturbances using laser sensors are described in the source [10]. The first one is the symmetrical estimation method (Fig. 3).

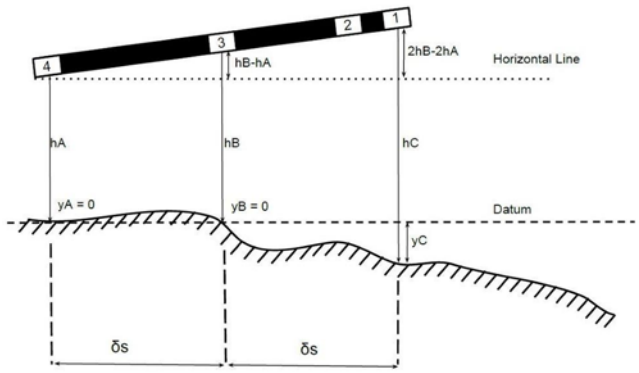


Fig. 3. The symmetrical road roughness estimation method [10]

This estimation method is more reliable when measuring the road roughness waves of 3 m and longer. The symmetrical method is similar to the previously mentioned Japanese method, however the laser sensors are arranged at the distance of $\sigma s = 0.6$ m m. In Fig. 3, the laser distance sensors are marked by numbers 1, 3, 4. After recording the data of the sensors, the roughness is calculated by using equations:

$$y_n = y_{n-1} + \sum_{i=1}^{n} U_i \tag{3}$$

$$U = -(h_1 - 2h_3 + h_4) \quad (4)$$

Here y and U indicate the average height at the distance of σs.

2.3. Asymmetrical estimation method

The asymmetrical estimation method is used as a supporting method for the symmetrical one. This method is designed for determining the smaller waves of the road roughness. The system also uses

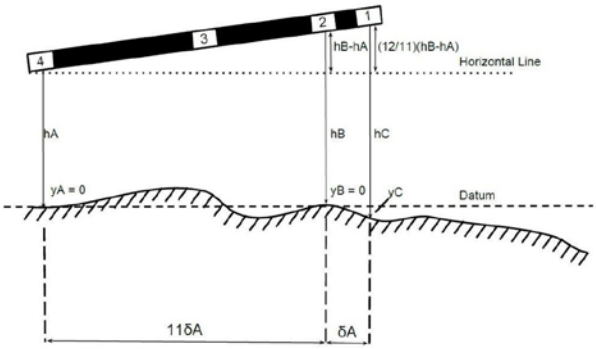


Fig. 4. The asymmetrical road roughness estimation method [10]

3 laser sensors, only they are arranged asymmetrically (Fig. 4).

The distance between the sensors 1 and 2 is m, the distance between the sensors 2 and 4 is m. The recorded values of each sensor are used when calculating the road roughness average by using the equations:

$$W = -(h_1 - \frac{12}{11}h_2 + \frac{1}{11}h_4) \quad (5)$$

$$y_n = \frac{12}{11}y_{n-1} - \frac{1}{11}y_{n-12} + W_n \quad (6)$$

Here y and W indicate the average height and measurement parameters at the distance of σA.

The presented road roughness estimation methods limit themselves to the processing of the signals measured by the laser distance to the road surface sensors, however when mounting the sensors on the car it is equally important to evaluate the inertia of the bodywork.

3. Experimental Research

Car Mercedes-Benz E350 BlueTec 4MATIC 2014 year is used for the research. The weight of the car during the research is 1956 kg, the overall dimensions – length 4879 mm, width 1853 mm, height 1474 mm, wheel base 2873 mm, researched with tires of dimensions 245/45 R17. Recording the dynamic parameters, the mobile *Kistler Group* measurement equipment was mounted on the car: three distance measurement sensors *Corrsys-Datron HF-500C* (Fig. 5), two (for separate axes) wheel vector sensors *Kistler RV-4*, an optical velocity and driver distance measurement device *Correvit S-350 Aqua*,

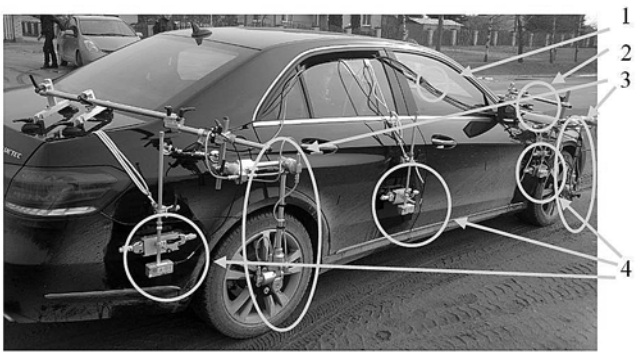


Fig. 5. A car with the mounted mobile Kistler Group measurement equipment

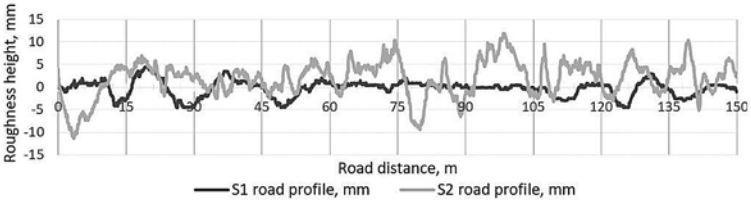


Fig. 6. The profiles of the road surfaces S1 and S2

Table 1. The technical parameters of the used equipment

Layout of the equipment (Fig 5)	Name and model of the device	Measured parameter, units	Measurement accuracy	Principle of the measurement	Limits of the measurement
1	Corrsys-Datron TANS-3	Body pitch rate, θ, deg	±0.1%	IMU	±150°/s
2	Correvit S-350 Aqua	Vehicle speed, v_x, km/h	±0.2%	Optical	0,5-250 km/h
3	*Kistler RV-4*	*Suspension deflection, Δz, mm*	±1 *mm*	*Mechanical*	±200 *mm*
4	Corrsys-Datron HF-500C	Distance, φ, m	±0.2%	Laser	125-625 mm

vehicle inertia measurement unit (IMU) *Corrsys-Datron TANS-3*. The technical parameters of the used equipment are presented in table 1.

All of the measured parameters were registered at the frequency of 200 Hz to Data Acquisition Systems *DAS-3*, and later processed on the computer by using the *TurboLab 6.0* software.

Two dry asphalt road sections of length of 150 m were chosen for the experimental research. The profiles of the two selected different road surfaces are presented in the Fig. 6. The largest deflection of the surface S1 is 4.8 mm, and of the surface S2 – 11.9 mm [26].

When carrying out the research, the car Mercedes-Benz E350 with the mounted equipment was driven at a steady speed of 80 km/h. When processing the research results, the calculations were done using 3 methods:

- The described Japanese method;
- The improved Japanese method – using only the values of h_{i-1} and h_{i+1}, and compensating the suspension displacements (deflections) of the moving car;
- The improved Japanese method – using only the values of h_{i-1} and h_{i+1}, and compensating the body pitch rate oscillations.

When improving the Japanese method, only the values of h_{i-1}, h_{i+1} with the compensation of the suspension displacements were selected. In other case, the compensation the body pitch rate oscillations was added. The body pitch rate oscillation is described in the "half-car" model (Fig. 7).

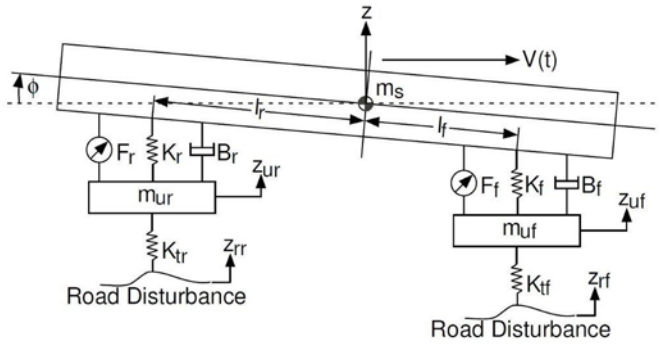


Fig. 7. "Half-car" model [23]

The "half-car" model consists of three bodies: m_s – car's bodywork (sprung mass), m_{ur} ir m_{uf} – the front and rear wheels with axles (unsprung masses). This model has 4 degrees of freedom: the vertical displacement of sprung mass z and transverse about the transverse axis ϕ, the displacements of the unsprung masses z_{ur} and z_{uf}. Independent excitations from the road are marked by displacements z_{rr} and z_{rf}, the front and rear axles' suspension forces F_r ir F_f.

The parameters in the "half-car" model are described by the Newton-Euler movement equations.

The vertical movement of the sprung mass:

$$m_s\ddot{z} = -(B_f + B_r)\dot{z} + (l_f B_f - l_r B_r)\dot{\phi} - K_f(z_{rf} - z_{uf}) + B_f\dot{z}_{uf} - K_r(z_{sr} - z_{ur}) + B_r\dot{z}_{ur} + F_f + F_r \tag{7}$$

The pivoting about the transverse axis:

$$I\ddot{\phi} = (l_f B_f - l_r B_r)\dot{z} - (l_f^2 B_f + l_r^2 B_r)\dot{\phi} + l_f K_f(z_{sf} - z_{uf}) - l_f B_f\dot{z}_{uf} - l_r K_r(z_{sr} - z_{ur}) + l_r B_r\dot{z}_{ur} - l_f F_f + l_r F_r \tag{8}$$

The movement of the front unsprung mass:

$$m_{uf}\ddot{z}_{uf} = -K_{tf}z + B_f\dot{z} + l_f K_{tf}\phi - l_f B_f\dot{\phi} + (K_f + K_{tf})(z_{sf} - z_{uf}) - B_f\dot{z}_{uf} - K_{tf}z_{rf} - F_f \tag{9}$$

The movement of the rear unsprung mass:

$$m_{ur}\ddot{z}_{ur} = -K_{tr}z + B_r\dot{z} - l_r K_{tr}\phi + (K_r + K_{tr})(z_{sr} - z_{ur}) - B_r\dot{z}_{ur} + K_{tr}z_{rr} - F_r \tag{10}$$

In the model, the values of the rigidity of the tire and the suspension – K_{tf}, K_{tr}, K_f ir K_r, as well as the values of the suspensions damping $B_f - B_r$ ir are used.

In this work, the compensation of the body pitch while improving the Japanese method is used. For this goal, the displacement of the bodywork to the front and rear axles, which is formed by the pivoting around the transverse axis, is calculated (Fig. 7):

$$\Delta z_r = l_r \cdot \tan\phi \tag{11}$$

$$\Delta z_f = l_f \cdot \tan\phi \tag{12}$$

Each Δ value is compensated from at the same time measured values of h_{i-1}, h_{i+1}.

4. Results and discussion

After processing the measured data using 3 described methods, the obtained results are presented in the Fig. 8. and Fig. 9. For better comparison of the methods, the results of one road section are shown in the same graph.

In order to evaluate the methods even more accurately, the validation based on the root mean squared error was used. The validation was done with the MATLAB software, and the results are presented in the Table 2. The method most accurately matches the road profile when the body pitch rate oscillation is compensated. Validating by using this method, the root mean squared error is the lowest in the section S1 – 0.5728, and in the section S2 – 2.4538. The Japanese method dies not fully reflect the road profile because, when doing the calculations using the 1 equation, the average of the measured heights

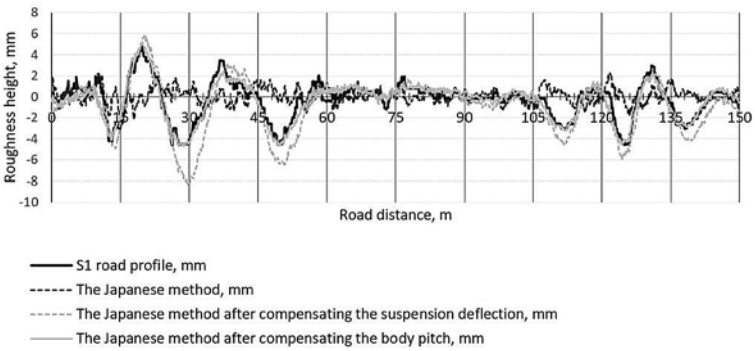


Fig. 8. The road profile determined by using three methods compared to the real road profile of the section S1.

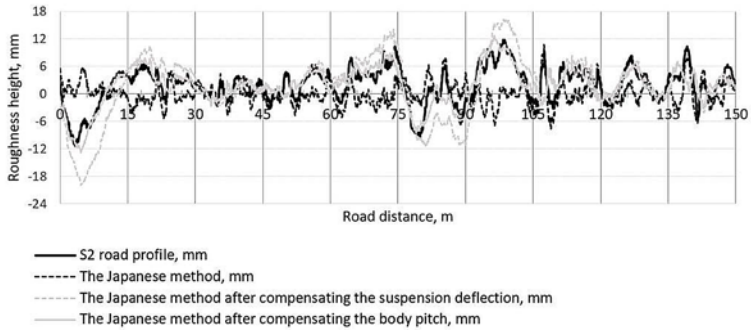


Fig. 9. The road profile determined by using three methods compared to the real road profile of the section S2.

Table 2. The results of the validation of the road roughness estimation methods.

Measurement method	Validation based on the Root Mean Squared Error	
	S1 section	S2 section
The Japanese method	1.9428	4.7157
The Japanese method after compensating the suspension deflection	1.2804	4.2886
The Japanese method after compensating the body pitch rate oscillations	0.5728	2.4538

and by the side sensors is subtracted from value of the height of the middle sensor.

The improved Japanese method – using only the h_{i-1} and h_{i-+} values and by compensating the displacement of the suspensions for the moving car is not a sufficiently accurate method to determine the road surface. Moving through the road disturbances the car oscillates and the suspension damps the oscillations of the bodywork, that is why the displacement of the suspension is compensated, which is larger than the actual road disturbance.

5. Conclusion

After reviewing the road roughness estimation methods using the laser distance sensors and carrying out the experimental research with the sensors mounted on a standard car, the following conclusions were made:

- the widely applied Japanese method for the practical estimation of the road roughness based on the data of the three laser distance sensors arranged at the distance of 1.5 m is not sufficient when the sensors are mounted on the bodywork of a standard car;
- when determining the road disturbances based on the data of the laser distance sensors mounted on a car, correction based on the suspension travel is not sufficient as it does not eliminate the oscillations of the bodywork;
- after applying the compensation of the body pitch rate oscillations, the best results of the road surface roughness was obtained, and, after validating it, based on the real road profile in road sections of different smoothness, the root mean squared errors that are lower from 1.92 to 3.39 times are obtained. After applying the less effective compensation of the suspensions travel, the root mean squared errors are lower only 1.1-1.52 times.

References

1. Ahmed A, Saeed T U, Labi S. Estimation of rest periods for newly constructed reconstructed pavements, Transport 2016; 31(2): 183–191, https://doi.org/10.3846/16484142.2016.1193050.
2. Aki A, Rojanaarpa T, Nakano K, Suda Y, Takasuka N, Isogai T, Kawai T. Road Surface Recognition Using Laser Radar for Automatic Platooning. IEEE Transaction on intelligent transportation systems 2016; 17(10): 2800–2810, http://dx.doi.org/10.1109/TITS.2016.2528892.
3. Becker C, Els S. Profiling of rough terrain. International Journal of Vehicle Design 2014; 64: 240–261, http://dx.doi.org/10.1504/IJVD.2014.058500.
4. Castillo Aguilar J J, Cabrera Carrillo J A, Guerra Fernández A J, Carabias Acosta E. Robust Road Condition Detection System Using In-Vehicle Standard Sensors, Sensors 2015; 15(12): 32056-32078, https://doi.org/10.3390/s151229908.
5. Doumiati M, Erhart S, Martinez J, Sename O, Dugard L. Adaptive control scheme for road profile estimation application to vehicle dynamics. The International Federation of Automatic Control 2014: 8445–8450, http://dx.doi.org/10.3182/20140824-6-ZA-1003.00986.
6. Doumiati M, Martinez J, Sename O, Dugard L, Lechner D. Road profile estimation using an adaptive Youla–Kučera parametric observer: Comparison to real profilers, Control Engineering Practice 2015: 1–9, http://dx.doi.org/10.1016/j.conengprac.2015.12.020.
7. El-Ashmawy K L A. A simple technique for road surface modelling, Geodesy and Cartography 2016; 42(3): 106–114, http://dx.doi.org/10.3846/20296991.2016.1226392.
8. Harris N K, Gonzalez A, O`Brien E J, McGetrick P. Characterisation of pavement profile heights using accelerometer readings and a combinatorial optimisation technique. Journal of Sound and Vibration 2010, 329: 497–508, http://dx.doi.org/10.1016/j.jsv.2009.09.035.
10. Kilic F, Hilsmann J. Application and Improvement of the TRRL (Transport and Road Research Laboratory) High-Speed Laser Profilometer Algorithm with Sensor Fusion. IFAC (International Federation of Automatic Control) Papers On Line 1966; 49-15: 260-265, http://dx.doi.org/10.1016/j.ifacol.2016.07.761.
11. Levulytė L, Žuraulis V, Sokolovskij E. The research of dynamic characteristics of the vehicle driving over road roughness. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(4): 518–525.
12. Martinez-Arroyo M, Enrique S L. Learning an Optimal Naive Bayes Classifier. The 18th International Conference on Pattern Recognition 2006, http://dx.doi.org/10.1109/ICPR.2006.748.
13. Múčka P. Current approaches to quantify the longitudinal road roughness, International Journal of Pavement Engineering 2016; 17(8): 659–679, https://doi.org/10.1080/10298436.2015.1011782.
14. Múčka P. International Roughness Index specifications around the world, Road Materials and Pavement Design 2016: 1-37, https://doi.org/10.1080/14680629.2016.1197144.
15. O`Brien, E J, McGetrick P, Gonzalez A. Identification of Road Irregularities via Vehicle Accelerations. Transport Research Arena Europe 2010.
16. Pakorski j, Reński A, Sar H. System for investigation of friction properties of the road surface, The Baltic Journal of Road and Bridge Engineering 2015; 10(2): 126–131, https://doi.org/10.3846/bjrbe.2015.16.
17. Raslavičius L, Pakalnis A, Keršys A, Skvireckas R, Juodvalkis D. Investigation of asphalt texture roughness on friction evolution for wheeled vehicles. Transport 2016; 31(2): 133–141, http://dx.doi.org/10.3846/16484142.2016.1189960.
18. Sahlholm P, Jansson H, Kozica E, Johansson K H. A sensor and data fusion algorithm for road grade estimation, 5th IFAC Symposium on Advances in Automotive Control 2007; 40(10): 55–62, http://dx.doi.org/10.3182/20070820-3-US-2918.00010.
19. Sayers M W. Two Quarter-car models for defining road roughness IRI ir HRI. Transportation Research Record 1989; 1215: 165–172.
20. Sayers M W, Karamihas S M. The Little Book of Profiling. Basic Information about Measuring and Interpreting Road Profiles. Michigan 1998.
21. Souza R O, Neto S D, Farias M M. Improving Pavements With Long-Term Pavement Performance: Products for Today and Tomorrow. Papers From the 2003–2004 International Contest on Long-Term Pavement Performance Data Analysis 2006: 5–6.
22. Tomiyama K, Kawamura A. Application of lifting wavelet transform for pavement surface monitoring by use of a mobile profilometer. International Journal of Pavement Research and Technology 2016; 9(5): 345–353, http://dx.doi.org/10.1016/j.ijprt.2016.08.007.
23. Vaughan J E. Active and Semi-Active Control to Counter Vehicle Payload Variation. A thesis, The George W. Woodruff School of Mechanical Engineering, George Institute of Technology 2004.

24. Yuan Z, Zhang X, Liu S, Han X, Du Y. Laser Line Recognition for Autonomous Road Roughness Measurement. The 5th Annual IEEE International conference on Cyber Technology in Automation, Control and Inteligent Systems 2015: 436–440, http://dx.doi.org/10.1109/CYBER.2015.7287977.
25. Žuraulis V, Levulytė L, Sokolovskij E. The Impact of Road Roughness on the Duration of Contact Between a Vehicle Wheel and Road Surface, Transport 2014; 29(4): 431–439, http://dx.doi.org/10.3846/16484142.2014.984330.
26. Žuraulis V, Pečeliūnas R, Jakutis G. Semi-active suspension influence on comfort sensation of a vehicle occupant. Agricultural Engineering, Research Papers 2014; 46(1): 116–124, https://doi.org/10.15544/ageng.2014.011.

Vytenis SURBLYS
Vidas ŽURAULIS
Edgar SOKOLOVSKIJ
Department of Automobile Transport
Vilnius Gediminas Technical University
Basanavičiaus str., 28 LT-03224 Vilnius, Lithuania

E-mails: vytenis.surblys@vgtu.lt, vidas.zuraulis@vgtu.lt, edgar.sokolovskij@vgtu.lt

Article citation info:
PAŚ J, ROSIŃSKI A. Selected issues regarding the reliability-operational assessment of electronic transport systems with regard to electromagnetic interference. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 375–381, http://dx.doi.org/10.17531/ein.2017.3.8.

Jacek PAŚ
Adam ROSIŃSKI

SELECTED ISSUES REGARDING THE RELIABILITY-OPERATIONAL ASSESSMENT OF ELECTRONIC TRANSPORT SYSTEMS WITH REGARD TO ELECTROMAGNETIC INTERFERENCE

WYBRANE ZAGADNIENIA OCENY NIEZAWODNOŚCIOWO-EKSPLOATACYJNEJ TRANSPORTOWYCH SYSTEMÓW ELEKTRONICZNYCH Z UWZGLĘDNIENIEM ZAKŁÓCEŃ ELEKTROMAGNETYCZNYCH*

The article presents issues associated with the reliability-operational analysis of electronic transport systems, which are operated in a defined environment. Intended or unintended (static or mobile) electromagnetic interference, which are present over a broad transport area, can cause interference of its functioning. That is why correct functioning of electronic transport systems in a distorted electromagnetic environment is so important. The articles measurement results of low frequency radiation basic characteristics for power supplies, which were carried out for two frequency ranges: (0÷400) Hz and (400÷6500) Hz. The impact of the load's voltage change on the electric field E [V/m] produced by the power supply and the impact of power supply load's current changes on the induction of the magnetic field B [nT] were presented. Next, a graph of relations in an electronic transport system, with regard to level of interference in electric and magnetic fields, was developed. This allowed to determine the relations, which enable to define the probability values of a system being in the distinguished states. The methodology of the reliability-operational analysis of electronic transport systems with regard to electromagnetic interference presented in this article, may be used during designing of electronic systems used in transport.

***Keywords**: operation, reliability, electromagnetic interference, electronic transport systems.*

W artykule przedstawiono zagadnienia związane z analizą niezawodnościowo-eksploatacyjną transportowych systemów elektronicznych, które są eksploatowane w określonym środowisku. Występujące na rozległym obszarze transportowym zaburzenia elektromagnetyczne zamierzone lub niezamierzone (stacjonarne lub ruchome) mogą być przyczyną zakłócenia ich funkcjonowania. Dlatego tak istotne jest prawidłowe funkcjonowanie transportowych systemów elektronicznych w zniekształconym środowisku elektromagnetycznym. W artykule przedstawiono wyniki pomiarów podstawowych charakterystyk promieniowania niskiej częstotliwości dla zasilaczy, które zostały przeprowadzone dla dwóch zakresów częstotliwości: (0÷400) Hz i (400÷6500) Hz. Zaprezentowano wpływ zmiany napięcia obciążenia na pole elektryczne E [V/m] wytwarzane przez zasilacz, oraz wpływ zmiany prądu obciążenia zasilacza na indukcję pola magnetycznego B [nT]. Następnie opracowano graf relacji w transportowym systemie elektronicznym z uwzględnieniem poziomów zakłóceń pola elektrycznego i magnetycznego. Umożliwiło to wyznaczenie zależności pozwalających na określenie wartości prawdopodobieństw przebywania systemu w wyróżnionych stanach. Zaprezentowana w artykule metodyka analizy niezawodnościowo-eksploatacyjnej transportowych systemów elektronicznych z uwzględnieniem zakłóceń elektromagnetycznych może być użyta podczas projektowania systemów elektronicznych stosowanych w transporcie.

***Słowa kluczowe**: eksploatacja, niezawodność, zakłócenia elektromagnetyczne, transportowe systemy elektroniczne.*

1. Introduction

Electronic transport systems function in different, often extreme, operational conditions. Many years of observations of their use process in a transport environment confirm the dependence of their correct operation on the reliability of the components and effective management of their operation process [6, 7, 8]. The analysis of operational phenomena should, therefore, take into account not only the reliability approach but also the effectiveness of operational management. For this purpose, the authors imitated the phenomena undergoing in reality (including electromagnetic interference) in a research model of an electronic transport system.

Electronic transport systems are elements in many transport systems. Their proper functioning significantly impacts the safety and efficiency of the process of transporting humans and cargo. Unreliability of electrical equipment and operator errors may lead to safety hazard states [11, 15, 18, 30]. The theory of unreliability deals with the analysis of the impact of equipment damage and operator errors on defined unreliability indicators [24, 26]. The scope of interest of the theory of safety are the results of damages and errors, which lead to safety hazards. Very important is the issue of correct determination which of the system's states can be deemed permissible or impermissible, from the safety point of view.

The safety hazard state may be a reversible state, when there is a possibility to take actions aimed at restoring the state of full worthiness (e.g. Diagnosing a damage and attempting repair, correcting an operator error, neutralising an external event). The

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

execution of a counter-action has to be performed in the availability time. In case of the counter-action being ineffective or is taking too long, the system switches from the safety hazard state to a hazard state, also called the safety unreliability state.

An increase of the level of electronic transport systems' safety may be executed through improving their reliability parameters. An increase in reliability can be achieved by the increased reliability of components or using redundant structures [20, 21, 25, 32]. The first solution is aimed at preventing damages. In the second case, the use of double or triple redundancy leads to, admittedly, expansion of the system but enables tolerating the occurring damages. Redundancy may concern the devices subassemblies, system modules, as well as e.g. computers controlling transport processes. The quality of information [10, 28, 29], which the systems receive from sensors is also important [4, 9]. Some scientific papers propose to use fuzzy logic [27] or artificial neuron networks [5]. Electronic transport systems are also significantly impacted by vibrations [2] but they are not taken into consideration in this article.

Electronic transport systems operate in various operating conditions present in transport areas. Their correct functioning is conditioned not only on the reliability of individual components comprising the system but also on the level of electromagnetic interference [12, 14] and the operational strategies adopted for implementation

The article presents measurements of the electromagnetic spectrum, which was created as a result of the use of power supplies used in electronic transport systems. An analysis of the obtained results made it possible to develop a research model of an electronic transport system and then, to carry out the reliability-operational analysis with regard to electromagnetic interference.

2. Electronic transport systems

Mutual coexistence of electronic transport systems and their correct functioning in the transport domain in a defined electromagnetic environment and the functioning without introducing unacceptable electromagnetic interference into that environment, can be defined as electromagnetic compatibility [1, 3, 13].

In the case of the impact of electromagnetic interference on electronic transport systems, we can distinguish four operating states for these systems:

- the system does not react to external and internal interference – the interference level is too low, permissible interference levels were not exceeded, the system remains in a given operating state in which it currently is,
- the devices comprising the control system automatically remove interference through used passive or active filters, screening, correct distribution or system solutions,
- the occurrence of interference with values not exceeding the permissible ones, causes the system to switch from the worthiness state to the partial worthiness states – restoration of the worthiness state requires taking specific actions,
- the occurrence of interference with values exceeding the permissible ones, damages the system, and thus, the transition from the worthiness to the unworthiness state.

Electronic transport systems are installed in stationary and mobile facilities, and are used in a defined electromagnetic environment. Natural electromagnetic environment is shaped by the phenomena that occur on the ground, but is seriously distorted on the transport area. One of the reasons for that is the presence of a high number of electromagnetic fields, radiating in intended or unintended manners. Each electric or electronic device, supplied with electric power, produces its own electromagnetic field, which is associated with its operation [22, 23]. Electronic devices and systems, supplied by feeders of different types, are used over the transport area. They should function properly, regardless of the present interference, which have effect on them.

Electromagnetic interference with a broad frequency band, generated by different devices used over the transport area simultaneously, at the same time impact the electronic transport systems. The interference signal spectrum results from the operating frequency of systems (devices), their harmonics and the intermodulation frequencies, which are the result of non-linearity, e.g. of electronic elements.

Modern electronic transport systems used, i.a., in transport, are expected to satisfy many requirements. They can include, i.a., miniaturization, limited electrical energy consumption, high reliability. Introducing these limitations may result in the fact that the level of useful signals from devices may be comparable to the level of interference generated by, e.g., stationary and mobile interference sources (e.g. base and radio stations, MV and LV lines, transformer stations, commonly use electrical devices). That is why, it is necessary to perform a reliability-operational analysis of the systems, with regard to electromagnetic interference.

3. Reliability-operational analysis of electronic transport systems with regard to electromagnetic interference

In order to measure the low frequency radiation characteristics of power feeders of electronic transport systems, a spectrum analyser, with an operating frequency of 0 Hz to 100 kHz, was used. The radiation spectrum in the low frequency range, and this was adopted for the device functional analysis, is divided into two sub-ranges: ELF and VLF. Observing the changes of individual spectrum components radiated by the power supplies, their operating state can be characterized.

The measurement of the low frequency radiation basic characteristics was carried out for two frequency ranges: (0÷400) Hz and (400÷6500) Hz. A measurement above the frequency of 6.5 kHz is of little usefulness from the point of view of the device's diagnostics, since particular components of the electromagnetic field radiation have very small values. Tables 1 and 2 present the impact of the load's voltage change on the electric field E [V/m] produced by the power supply and the impact of power supply load's current changes on the induction of the magnetic field B [nT]. The E, B components

Table 1. Electrical field E [V/m] produced by tested power supplies for individual signal spectrum components (load current constant, change of the output voltage for a given power supply type)

Power supply type	P339		BS525		5353	
Frequency [Hz]	U=10V I=3A	U=20V I=3A	U=15V I=1A	U=30V I=1A	U=20V I=1,8A	U=48V I=1,8A
50	1,603	1,799	1,71	2,64	1,628	2,277
100	0,0053	0,0058	0,0062	0,0073	0,003	0,003
150	0,174	0,241	0,174	0,34	0,16	0,21
350	0,13	0,18	0,1294	0,14	0,12	0,22
1250	0,173	0,18	0,0196	0,021	0,016	0,015

Table 2. Induction of the magnetic field B [nT] produced by tested power supplies for individual signal spectrum components (load voltage constant, change of the output current for a given power supply type)

Power supply type	P339		BS525		5353	
Frequency [Hz]	U=10V I=1A	U=10V I=4A	U=30V I=0,5A	U=30V I=1A	U=48V I=0,5A	U=48V I=1,8A
50	1834,25	1883,65	3556,3	4358,6	2540,9	2851,02
100	527,84	901,57	53	109,5	44,8	125,9
150	5134,52	4623,81	188	590	512,9	575,4
250	845,28	1096,48	169	199,8	46,1	316,3
350	134,74	121,62	42,5	467,4	14,2	50,2
550	4,47	13,96	1,7	4,6	2,99	5,7

of the electromagnetic field are produced in an unintended way by the tested power supplies of the electronic transport systems. In order to determine the impact of the am parameters (load current, output voltage) of the system's power supply, a measurement stand for measuring diffused fields E, B produced by the tested devices was developed and constructed.

When testing the impact of load voltage on the spectrum of an electromagnetic field, the increase of signal spectrum harmonics component amplitude can be noticed for particular power supply types (tab. 1). For the P339 power supply, the biggest impact was noticed for the 150 Hz harmonics – the value of 0.174 V/m increased to 0.241 V/m, with the increase of the supply voltage from 10V to 20V. Also for the power grid's component frequency of 50Hz there was an increase of amplitude, but it was smaller and amounted to only 10.89%. For the remaining power supplies, an increase of the amplitude of individual harmonics, together with an increase of the supply voltage, can also be noticed. The biggest increase was for the frequency of 100 Hz, for the BS525 power supply and amounted to 48.8%. For higher frequencies (above 350 Hz), individual field components with the change of the supply load voltage, did not have high amplitude values. The increase of the amplitude values was very small, together with the increase of the load voltage.

A similar phenomenon of the increase in the amplitude of electromagnetic field's individual harmonics can be observed for magnetic field B induction. An increase of the field's induction was noticed during the increase of the load current drawn from a power supply (tab. 2). A significant increase of the harmonics amplitude of the electromagnetic field can be observed for higher frequencies, i.e., 250, 350, 550 Hz. For the frequency of 350 Hz, the increase of the harmonics amplitude for the BS525 power supply was as much as 90.9%. However, with the increase of the load current, increases of individual negative harmonics were also observed. This phenomenon, i.a., can be associated with returning the field's energy from some spectrum bands to others. By thoroughly studying the increase of magnetic field's B induction and electric field's E current, we can obtain information about the state of the power supply load's current and voltage.

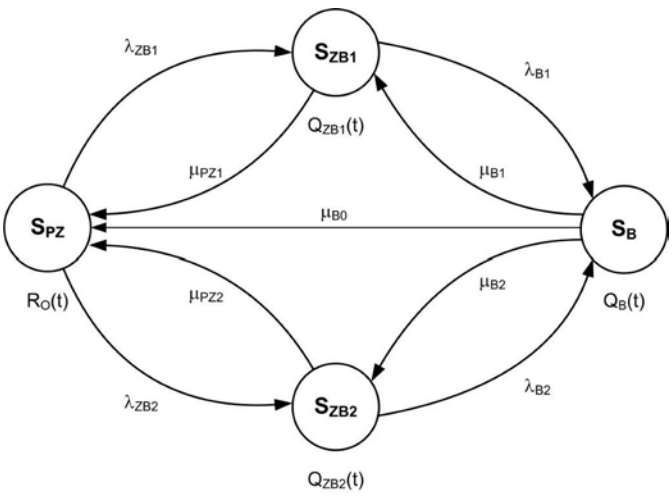


Fig. 1. Relations in an electronic transport system

Analysing the obtained results, we can conclude that an electronic transport system (operated in a transport environment, with electromagnetic interference present) can be in the distinguished states [19, 31]. Therefore, the idea to carry out the reliability-operational analysis seems reasonable. Thanks to that, we will obtain relations, which allow to determine the probability values of the system being in the distinguished operational states

By conducting the reliability and exploitation analysis of the electronic transport system, it is possible to illustrate the relationships in it, as shown in Figure 1.

Denotations in figures:

$R_O(t)$ – the function of probability of system staying in state of full ability S_{PZ},

$Q_{ZB1}(t)$–the function of probability of system staying in state of the impendency over safety S_{ZB1},

$Q_{ZB2}(t)$–the function of probability of system staying in state of the impendency over safety S_{ZB2},

$Q_B(t)$ – the function of probability of system staying in state of unreliability of safety S_B,

λ_{ZB1} – transition rate from the state of full ability S_{PZ} into the state of the impendency over safety S_{ZB1},

λ_{ZB2} – transition rate from the state of full ability S_{PZ} into the state of the impendency over safety S_{ZB2},

μ_{PZ1} – transition rate from the state of the impendency over safety S_{ZB1} into the state of full ability S_{PZ},

μ_{PZ2} – transition rate from the state of the impendency over safety S_{ZB2} into the state of full ability S_{PZ},

λ_{B1} – transition rate from the state of the impendency over safety S_{ZB1} into the state of unreliability of safety S_B,

λ_{B2} – transition rate from the state of the impendency over safety S_{ZB2} into the state of unreliability of safety S_B,

μ_{B1} – transition rate from the state of unreliability of safety S_B into the state of the impendency over safety S_{ZB1},

μ_{B2} – transition rate from the state of unreliability of safety S_B into the state of the impendency over safety S_{ZB2},

μ_{B0} – transition rate from the state of unreliability of safety S_B into the state of full ability S_{PZ}.

Full worthiness state S_{PZ} is a state, in which the electronic transport system functions correctly. The safety hazard state S_{ZB1} is a state, in which the electronic transport system is partially fit for use (interference of the electric field do not exceed permissible values) The safety hazard state S_{ZB2} is a state, in which an electronic

transport system is partially fit for use (interference of the magnetic field do not exceed permissible values) The safety hazard state S_B is a state, in which the electronic transport system is unfit for use (interference of the electric and magnetic fields exceed permissible values).

If an electronic transport system is in a state of full worthiness S_{PZ} and there appears an electric field interference, then it switches to the safety hazard state S_{ZB1} with the intensity λ_{ZB1}. If the system is in the safety hazard state S_{ZB1}, then it is possible to switch to the full worthiness state S_{PZ}, provided that actions are taken aimed at restoring the worthiness state.

In case of a safety hazard state S_{ZB1} and additionally, there appears an interference of the magnetic field, then the system switches to the safety unreliability state S_B with the intensity λ_{B1}. Return transition of the safety hazard state S_{ZB1} from the safety unreliability state S_B is possible, provided actions are taken, aimed at mitigating the level of magnetic field interference, to a value in the permissible range.

If an electronic transport system is in a state of full worthiness S_{PZ} and there appears a magnetic field interference, then it switches to the safety hazard state S_{ZB2} with the intensity λ_{ZB2}. If the system is in the safety hazard state S_{ZB2}, then it is possible to switch to the full worthiness state S_{PZ}, provided that actions are taken aimed at restoring the worthiness state.

In case of a safety hazard state S_{ZB2} and additionally, there appears an interference of the electric field, then the system switches to the safety unreliability state S_B with the intensity λ_{B2}. Return transition of the safety hazard state S_{ZB2} from the safety unreliability state S_B is possible, provided actions are taken, aimed at mitigating the level of electric field interference, to a value in the permissible range.

If an electronic transport system is in the safety unreliability state S_B and there are interference in the magnetic and electric fields, then the system switches to the full worthiness state S_{PZ} with the intensity μ_{B0}.

The system illustrated in fig. 1 may be described by the following Chapman–Kolmogorov equations:

$$
\begin{aligned}
R_0'(t) &= -\lambda_{ZB1}\times R_0(t) + \mu_{PZ1}\times Q_{ZB1}(t) - \lambda_{ZB2}\times R_0(t) + \mu_{PZ2}\times Q_{ZB2}(t) + \mu_{B0}\times Q_B(t)\\
Q_{ZB1}'(t) &= \lambda_{ZB1}\times R_0(t) - \mu_{PZ1}\times Q_{ZB1}(t) - \lambda_{B1}\times Q_{ZB1}(t) + \mu_{B1}\times Q_B(t)\\
Q_{ZB2}'(t) &= \lambda_{ZB2}\times R_0(t) - \mu_{PZ2}\times Q_{ZB2}(t) - \lambda_{B2}\times Q_{ZB2}(t) + \mu_{B2}\times Q_B(t)\\
Q_B'(t) &= \lambda_{B1}\times Q_{ZB1}(t) + \lambda_{B2}\times Q_{ZB2}(t) - \mu_{B0}\times Q_B(t) - \mu_{B1}\times Q_B(t) - \mu_{B2}\times Q_B(t)
\end{aligned}
\tag{1}
$$

Given the initial conditions:

$$
\begin{aligned}
&R_0(0) = 1\\
&Q_{ZB1}(0) = Q_{ZB2}(0) = Q_B(0) = 0
\end{aligned}
\tag{2}
$$

Laplace transform yields the following system of linear equations:

$$
\begin{aligned}
s\times R_0^*(s) - 1 &= -\lambda_{ZB1}\times R_0^*(s) + \mu_{PZ1}\times Q_{ZB1}^*(s) - \lambda_{ZB2}\times R_0^*(s) + \mu_{PZ2}\times Q_{ZB2}^*(s) + \mu_{B0}\times\\
s\times Q_{ZB1}^*(s) &= \lambda_{ZB1}\times R_0^*(s) - \mu_{PZ1}\times Q_{ZB1}^*(s) - \lambda_{B1}\times Q_{ZB1}^*(s) + \mu_{B1}\times Q_B^*(s)\\
s\times Q_{ZB2}^*(s) &= \lambda_{ZB2}\times R_0^*(s) - \mu_{PZ2}\times Q_{ZB2}^*(s) - \lambda_{B2}\times Q_{ZB2}^*(s) + \mu_{B2}\times Q_B^*(s)\\
s\times Q_B^*(s) &= \lambda_{B1}\times Q_{ZB1}^*(s) + \lambda_{B2}\times Q_{ZB2}^*(s) - \mu_{B0}\times Q_B^*(s) - \mu_{B1}\times Q_B^*(s) - \mu_{B2}\times Q_B^*(s)
\end{aligned}
\tag{3}
$$

Probabilities of system staying in a distinguished functional states in symbolic (Laplace) terms have the following form:

$$
\begin{aligned}
R_0^*(s) &= \frac{b_2\times\lambda_{B1}\times\mu_{B1} - b_1\times b_2\times c + b_1\times\lambda_{B2}\times\mu_{B2}}{\begin{aligned}&a\times b_2\times\lambda_{B1}\times\mu_{B1} - a\times b_1\times b_2\times c + a\times b_1\times\lambda_{B2}\times\mu_{B2} + b_2\times c\times\lambda_{ZB1}\times\mu_{PZ1} + b_1\times c\times\lambda_{ZB2}\times\mu_{PZ2} +\\ &+b_2\times\lambda_{B1}\times\mu_{B0}\times\lambda_{ZB1} + b_1\times\mu_{B0}\times\lambda_{B2}\times\lambda_{ZB2} - \lambda_{B1}\times\mu_{B1}\times\lambda_{ZB2}\times\mu_{PZ2} + \lambda_{B1}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ2} +\\ &+\lambda_{B2}\times\mu_{B1}\times\mu_{PZ1}\times\lambda_{ZB2} - \lambda_{B2}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ1}\end{aligned}}\\
Q_{ZB1}^*(s) &= -\frac{b_2\times c\times\lambda_{ZB1} + \lambda_{B2}\times\mu_{B1}\times\lambda_{ZB2} - \lambda_{B2}\times\mu_{B2}\times\lambda_{ZB1}}{\begin{aligned}&a\times b_2\times\lambda_{B1}\times\mu_{B1} - a\times b_1\times b_2\times c + a\times b_1\times\lambda_{B2}\times\mu_{B2} + b_2\times c\times\lambda_{ZB1}\times\mu_{PZ1} + b_1\times c\times\lambda_{ZB2}\times\mu_{PZ2} +\\ &+b_2\times\lambda_{B1}\times\mu_{B0}\times\lambda_{ZB1} + b_1\times\mu_{B0}\times\lambda_{B2}\times\lambda_{ZB2} - \lambda_{B1}\times\mu_{B1}\times\lambda_{ZB2}\times\mu_{PZ2} + \lambda_{B1}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ2} +\\ &+\lambda_{B2}\times\mu_{B1}\times\mu_{PZ1}\times\lambda_{ZB2} - \lambda_{B2}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ1}\end{aligned}}\\
Q_{ZB2}^*(s) &= -\frac{b_1\times c\times\lambda_{ZB2} + \lambda_{B1}\times\mu_{B2}\times\lambda_{ZB1} - \lambda_{B1}\times\mu_{B1}\times\lambda_{ZB2}}{\begin{aligned}&a\times b_2\times\lambda_{B1}\times\mu_{B1} - a\times b_1\times b_2\times c + a\times b_1\times\lambda_{B2}\times\mu_{B2} + b_2\times c\times\lambda_{ZB1}\times\mu_{PZ1} + b_1\times c\times\lambda_{ZB2}\times\mu_{PZ2} +\\ &+b_2\times\lambda_{B1}\times\mu_{B0}\times\lambda_{ZB1} + b_1\times\mu_{B0}\times\lambda_{B2}\times\lambda_{ZB2} - \lambda_{B1}\times\mu_{B1}\times\lambda_{ZB2}\times\mu_{PZ2} + \lambda_{B1}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ2} +\\ &+\lambda_{B2}\times\mu_{B1}\times\mu_{PZ1}\times\lambda_{ZB2} - \lambda_{B2}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ1}\end{aligned}}\\
Q_B^*(s) &= -\frac{b_2\times\lambda_{B1}\times\lambda_{ZB1} + b_1\times\lambda_{B2}\times\lambda_{ZB2}}{\begin{aligned}&a\times b_2\times\lambda_{B1}\times\mu_{B1} - a\times b_1\times b_2\times c + a\times b_1\times\lambda_{B2}\times\mu_{B2} + b_2\times c\times\lambda_{ZB1}\times\mu_{PZ1} + b_1\times c\times\lambda_{ZB2}\times\mu_{PZ2} +\\ &+b_2\times\lambda_{B1}\times\mu_{B0}\times\lambda_{ZB1} + b_1\times\mu_{B0}\times\lambda_{B2}\times\lambda_{ZB2} - \lambda_{B1}\times\mu_{B1}\times\lambda_{ZB2}\times\mu_{PZ2} + \lambda_{B1}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ2} +\\ &+\lambda_{B2}\times\mu_{B1}\times\mu_{PZ1}\times\lambda_{ZB2} - \lambda_{B2}\times\mu_{B2}\times\lambda_{ZB1}\times\mu_{PZ1}\end{aligned}}
\end{aligned}
\tag{4}
$$

where:

$$a = s + \lambda_{ZB1} + \lambda_{ZB2}$$
$$b_1 = s + \mu_{PZ1} + \lambda_{B1}$$
$$b_2 = s + \mu_{PZ2} + \lambda_{B2}$$
$$c = s + \mu_{B1} + \mu_{B2} + \mu_{B0}$$

Solution to the above set of equations in the time domain is the next step in the analysis and is not discussed here.

4. Modelling an operation process of an electronic transport system

Computer simulation and computer-aided analysis facilitate to relatively quickly determine the influence of change in reliability-exploitation parameters of individual components on reliability of the entire system. Of course, the reliability structure of both the entire system and its components has to be known beforehand.

Using computer aided allows to perform the calculation of the value of probability of system staying in state of full ability R_O. That procedure is illustrated with below example.

Example

The following quantities were defined for the system:

– test duration - 1 year (values of this parameter is given in [h]):

$$t = 8760\ [h]$$

– non-damageability of the system in the case electric field interference, with values in the permissible range:

$$R_{ZB1}(t) = 0{,}949$$

– non-damageability of the system in the case magnetic field interference, with values in the permissible range:

$$R_{ZB2}(t) = 0{,}991$$

– transition rate from the state of the impendency over safety S_{ZB1} into the state of unreliability of safety S_B:

$$\lambda_{B1} = 0{,}000001$$

– transition rate from the state of the impendency over safety S_{ZB2} into the state of unreliability of safety S_B:

$$\lambda_{B2} = 0{,}000006$$

– transition rate from the state of unreliability of safety S_B into the state of full ability S_{PZ}:

$$\mu_{B0} = 0{,}01$$

– transition rate from the state of unreliability of safety S_B into the state of the impendency over safety S_{ZB1}:

$$\mu_{B1} = 0{,}1$$

– transition rate from the state of unreliability of safety S_B into the state of the impendency over safety S_{ZB2}:

$$\mu_{B2} = 0{,}1$$

Knowing the value of reliability $R_{ZB1}(t)$, transition rate from the state of full ability into the state of the impendency over safety S_{ZB1} may be estimated. Provided the up time is described by exponential distribution, the following relationship can be used:

$$R_{ZB1}(t) = e^{-\lambda_{ZB1} t} \text{ for } t \geq 0 \tag{5}$$

thus:

$$\lambda_{ZB1} = -\frac{\ln R_{ZB1}(t)}{t} \tag{6}$$

For $t = 8760\ [h]$ and $R_{ZB1}(t) = 0{,}949$ we obtain:

$$\lambda_{ZB1} = -\frac{\ln R_{ZB1}(t)}{t} = -\frac{\ln 0{,}949}{8760} = 0{,}000006\left[\frac{1}{h}\right] \tag{7}$$

Knowing the value of reliability $R_{ZB2}(t)$, transition rate from the state of full ability into the state of the impendency over safety S_{ZB2} may be estimated. Provided the up time is described by exponential distribution, the following relationship can be used:

$$R_{ZB2}(t) = e^{-\lambda_{ZB2} t} \text{ for } t \geq 0 \tag{8}$$

thus:

$$\lambda_{ZB2} = -\frac{\ln R_{ZB2}(t)}{t} \tag{9}$$

For $t = 8760\ [h]$ and $R_{ZB2}(t) = 0{,}991$ we obtain:

$$\lambda_{ZB2} = -\frac{\ln R_{ZB2}(t)}{t} = -\frac{\ln 0{,}991}{8760} = 0{,}000001\left[\frac{1}{h}\right] \tag{10}$$

For above initial values, by use of (4) equations, following results are obtained:

$$R_0^*(s) = \frac{\begin{array}{l} 3{,}85003\cdot 10^{11}\cdot s + 3{,}3\cdot 10^{11}\cdot \mu_{PZ1} + 5{,}5\cdot 10^{10}\cdot \mu_{PZ2} + 5\cdot 10^{17}\cdot s^2\cdot \mu_{PZ1} + 5\cdot 10^{17}\cdot s^2\cdot \mu_{PZ2} + \\ +1{,}050035\cdot 10^{17}\cdot s^2 + 5\cdot 10^{17}\cdot s^3 + 1{,}05003\cdot 10^{17}\cdot s\cdot \mu_{PZ1} + 1{,}050005\cdot 10^{17}\cdot s\cdot \mu_{PZ2} + \\ +1{,}05\cdot 10^{17}\cdot \mu_{PZ1}\cdot \mu_{PZ2} + 5\cdot 10^{17}\cdot s\cdot \mu_{PZ1}\cdot \mu_{PZ2} + 30000 \end{array}}{\begin{array}{l} 2{,}665021\cdot 10^{6}\cdot s + 1{,}050035\cdot 10^{17}\cdot s^2\cdot \mu_{PZ1} + 1{,}050035\cdot 10^{17}\cdot s^2\cdot \mu_{PZ2} - 5\cdot 10^{17}\cdot s^3\cdot \mu_{PZ1} + \\ +5\cdot 10^{17}\cdot s^3\cdot \mu_{PZ2} + 1{,}1200275\cdot 10^{12}\cdot s^2 + 1{,}05007\cdot 10^{17}\cdot s^3 + 5\cdot 10^{17}\cdot s^4 + \\ +4{,}35003\cdot 10^{11}\cdot s\cdot \mu_{PZ1} + 6{,}85003\cdot 10^{11}\cdot s\cdot \mu_{PZ2} + 1{,}05\cdot 10^{17}\cdot s\cdot \mu_{PZ1}\cdot \mu_{PZ2} + \\ +5\cdot 10^{17}\cdot s^2\cdot \mu_{PZ1}\cdot \mu_{PZ2} \end{array}} \tag{11}$$

Assuming $\mu_{PZ1} = 0{,}1$, $\mu_{PZ2} = 0{,}2$ and using the Laplace'a transformation we receive:

$$R_0(t) = 0{,}00000496942\cdot e^{-0{,}199947319\cdot t} + 0{,}000059996\cdot e^{-0{,}100006\cdot t} + 3{,}03804695\cdot e^{-0{,}21006058\cdot t} + \\ +0{,}9999350039 \tag{12}$$

Finally, we obtain:

$$R_O = 0{,}999935$$

The presented reliability-operational analysis of an electronic transport system, taking into account electromagnetic interference, allows numerical assessment of different types of solutions (technical and organizational), which can be implemented in order to mitigate the impact of electromagnetic interference on the system's functioning.

In order to mitigate the impact of electromagnetic interference on electronic transport systems, we need to determine: the interference source, interference receiver and the manner of the source coupling with the receiver. There are three main ways to limit the manners, the interference spreads:

– interference may be suppressed at the source (e.g. use of protective screening),
– execution of an electronic transport system, which shall be insensitive to the interference impacting it, which are present in an electromagnetic environment (use of electronic elements executed with proper technology),
– minimizing the transfer of interference through coupling channels (e.g. decoupling filters, groundings, gate drives [16, 17], screening, etc.).

In conclusion, if the constructors know the conditions of the electromagnetic environment (e.g. through the above presented measurements and reliability-operational analyses), in which the electronic transport system will probably be functioning, then the EMC requirements need to be met, which are usually known and can be taken into account when developing the device's structure.

5. Conclusion

Widespread use of electric and electronic systems in electronic transport systems causes the need of them functioning in a variety of systems, while being located very close to each other. This may result in an increase of the probability of interference in the systems' functioning, thus, being in states of partial worthiness. That is why, when designing electronic transport systems, they need to be prepared for operation in real conditions, meaning, being surrounded by other devices. The presented methodology of reliability-operational analysis of electronic transport systems, taking into account electromagnetic interference, may be helpful in this case. It allows to specify with numbers the probability values of the system being in the distinguished states.

In further studies, the authors plan to differentiate the sates of partial worthiness and highlight their subordinate states. This will allow more thorough mapping of the function of an electronic transport system utilized in an electromagnetic environment.

References

1. Billinton R, Allan RN. Reliability evaluation of power systems. New York: Plenum Press, 1996, https://doi.org/10.1007/978-1-4899-1860-4.
2. Burdzik R, Konieczny Ł, Figlus T. Concept of on-board comfort vibration monitoring system for vehicles. In the monograph Activities of Transport Telematics, editors: Mikulski J., TST 2013, CCIS 395. Heidelberg: Springer, 2013: 418-425, https://doi.org/10.1007/978-3-642-41647-7_51.
3. Charoy A. Interference in electronic equipment. Warsaw: WNT, 1999.
4. Dabrowski T, Bednarek M, Fokow K, Wisnios M. The method of threshold-comparative diagnosing insensitive on disturbances of diagnostic signals. Przeglad Elektrotechniczny - Electrical Review 2012; 88(11A):93-97.
5. Duer S, Zajkowski K, Duer R, Paś J. Designing of an effective structure of system for the maintenance of a technical object with the using information from an artificial neural network. Neural Computing & Applications 2012; 23(3): 913-925, https://doi.org/10.1007/s00521-012-1016-0.
6. Dyduch J, Paś J, Rosiński A. The basic of the exploitation of transport electronic systems. Radom: Publishing House of Radom University of Technology, 2011.
7. Garmabaki A.H.S., Ahmadi A., Mahmood Y.A., Barabadi A. Reliability modelling of multiple repairable units. Quality and Reliability Engineering International 2016; 32(7): 2329-2343, https://doi.org/10.1002/qre.1938.
8. Jacyna-Gołda I. Evaluation of operational reliability of the supply chain in terms of the control and management of logistics processes. In Safety and Reliability: Methodology and Applications - Proceedings of the European Safety and Reliability Conference ESREL 2014, editors: Nowakowski T., Młyńczak M., Jodejko-PietruczukA. &Werbińska-Wojciechowska S. London: CRC Press/Balkema, 2015: 549-558.
9. Kasprzyk Z, Rychlicki M. Analysis of phiysical layer model of WLAN 802.11g data transmission protocol in wireles networks used by telematic systems. In: Proceedings of the Ninth International Conference Dependability and Complex Systems DepCoS-RELCOMEX, given as the monographic publishing series - Advances in intelligent systems and computing", vol. 286. Springer, 2014: 265-274.
10. Lewiński A, Perzyński T, Toruń A. The analysis of open transmission standards in railway control and management. Communications in Computer and Information Science, vo. 329. Berlin Heidelberg: Springer-Verlag, 2012: 10-17.
11. Łubkowski P, Laskowski D. Selected issues of reliable identification of object in transport systems using video monitoring services. Communication in Computer and Information Science, vol. 471. Berlin Heidelberg, Springer, 2015: 59-68.
12. Ogunsola A, Mariscotti A. Electromagnetic compatibility in railways. Analysis and management. Springer-Verlag, 2013, https://doi.org/10.1007/978-3-642-30281-7.
13. Ott HW. Electromagnetic compatibility engineering. Wiley, 2009, https://doi.org/10.1002/9780470508510.
14. Paś J, Duer S. Determination of the impact indicators of electromagnetic interferences on computer information systems. Neural Computing & Applications 2012; 23(7): 2143-2157, https://doi.org/10.1007/s00521-012-1165-1.
15. Paś J. Operation of electronic transportation systems. Radom: Publishing House University of Technology and Humanities, 2015.
16. Perlicki K. Evaluation of the spatial distribution of birefringence in an optical-fiber link. Microwave and Optical Technology Letters 2004; 42 (2): 147-149, https://doi.org/10.1002/mop.20234.
17. Perlicki K. Impact of an alien wavelength on wavelength division multiplexing transmission quality. Photonics Letters of Poland 2012; 4(3): 118-120, https://doi.org/10.4302/plp.2012.3.12.
18. Perzyński T, Lewiński A, Łukasik Z. Safety analysis of accidents call system especially related to in-land water transport based on new telematic solutions. Communications in Computer and Information Science, Tools of Transport Telematics, vol 531. Berlin Heidelberg: Springer-Verlag, 2015: 90-98
19. Rosiński A. Modelling the maintenance process of transport telematics systems. Warsaw: Publishing House Warsaw University of Technology, 2015.
20. Siergiejczyk M, Krzykowska K, Rosiński A. Reliability assessment of cooperation and replacement of surveillance systems in air traffic. In: Proceedings of the Ninth International Conference Dependability and Complex Systems DepCoS-RELCOMEX, editors: W. Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak, J. Kacprzyk, given as the monographic publishing series - Advances in intelligent systems and computing", vol. 286. Springer, 2014: 403-411, https://doi.org/10.1007/978-3-319-07013-1_39.
21. Siergiejczyk M, Krzykowska K, Rosiński A. Reliability assessment of integrated airport surface surveillance system. In: Proceedings of the Tenth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX, editors: W.Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak, J. Kacprzyk, given as the monographic publishing series - Advances in intelligent systems and computing", vol. 365. Springer, 2015: 435-443, https://doi.org/10.1007/978-3-319-19216-1_41.
22. Siergiejczyk M, Paś J, Rosiński A. Train call recorder and electromagnetic interference. Diagnostyka 2015; 16(1): 19-22.
23. Siergiejczyk M, Paś J., Rosiński A. Issue of reliability-exploitation evaluation of electronic transport systems used in the railway environment with consideration of electromagnetic interference. IET Intelligent Transport Systems 2016; 10(9): 587-593, https://doi.org/10.1049/iet-its.2015.0183.
24. Siergiejczyk M, Paś J., Rosiński A. Modeling of process of exploitation of transport telematics systems with regard to electromagnetic interferences. In Tools of Transport Telematics, editors: Mikulski J., given as the monographic publishing series - Communications in Computer and Information Science", vol. 531. Berlin Heidelberg: Springer-Verlag, 2015: 99-107.
25. Siergiejczyk M, Rosiński A, Krzykowska K. Reliability assessment of supporting satellite system EGNOS. In the monograph New results in dependability and computer systems, editors: W. Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak, https://doi.org/10.1007/978-3-319-00945-2_32.
26. Siergiejczyk M, Rosiński A. Analysis of power supply maintenance in transport telematics system. Solid State Phenomena 2014; 210: 14-19, https://doi.org/10.4028/www.scientific.net/SSP.210.14.
27. Skorupski J, Uchroński P. A fuzzy reasoning system for evaluating the efficiency of cabin luggage screening at airports. Transportation Research Part C - Emerging Technologies 2015; 54: 157-175, https://doi.org/10.1016/j.trc.2015.03.017.
28. Stawowy M, Dziula P. Comparison of uncertainty multilayer models of impact of teleinformation devices reliability on information quality.

In: Proceedings of the European Safety and Reliability Conference ESREL 2015, editors: Podofillini L., Sudret B., Stojadinovic B., Zio E., Kröger W. CRC Press/Balkema, 2015: 2685-2691, https://doi.org/10.1201/b19094-351.
29. Stawowy M. Model for information quality determination of teleinformation systems of transport. In: Proceedings of the European Safety and Reliability Conference ESREL 2014, editors: Nowakowski T., Młyńczak M., Jodejko-Pietruczuk A., Werbińska-Wojciechowska S. CRC Press/Balkema, 2015: 1909-1914.
30. Sumiła M., Miszkiewicz A. Analysis of the problem of interference of the public network operators to GSM-R. In Tools of Transport Telematics, editors: Mikulski J., given as the monographic publishing series - Communications in Computer and Information Science", vol. 531. Berlin Heidelberg: Springer-Verlag, 2015: 76-82, https://doi.org/10.1007/978-3-319-24577-5_25.
31. Verma AK, Ajit S, Karanki DR. Reliability and safety engineering. London: Springer, 2010, https://doi.org/10.1007/978-1-84996-232-2.
32. Weintrit A., Dziula P., Siergiejczyk M., Rosiński A. Reliability and exploitation analysis of navigational system consisting of ECDIS and ECDIS back-up systems. The monograph Activities in Navigation - Marine Navigation And Safety Of Sea Transportation, editors: Weintrit A. London: CRC Press/Balkema, 2015: 109-115, https://doi.org/10.1201/b18513-17.

Jacek PAŚ
Military University of Technology
Faculty of Electronics
Institute of Electronic Systems
Division of Electronic Systems Exploitations
gen. S. Kaliskiego 2, 00-908 Warsaw, Poland

Adam ROSIŃSKI
Warsaw University of Technology
Faculty of Transport
Department of Telecommunications in Transport
Koszykowa 75, 00-662 Warsaw, Poland

E-mails: jacek.pas@wat.edu.pl , adro@wt.pw.edu.pl

Article citation info:
JACYNA-GOŁDA I, LEWCZUK K. The method of estimating dependability of supply chain elements on the base of technical and organizational redundancy of process. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 382–392, http://dx.doi.org/10.17531/ein.2017.3.9.

Ilona JACYNA-GOŁDA
Konrad LEWCZUK

THE METHOD OF ESTIMATING DEPENDABILITY OF SUPPLY CHAIN ELEMENTS ON THE BASE OF TECHNICAL AND ORGANIZATIONAL REDUNDANCY OF PROCESS

METODA SZACOWANIA NIEZAWODNOŚCI ELEMENTÓW ŁAŃCUCHA DOSTAW NA PODSTAWIE CHARAKTERYSTYK NADMIAROWOŚCI TECHNICZNEJ I ORGANIZACYJNEJ PROCESU

The quality of supply chain operation depends on quality of its particular elements, including warehouses. The paper presents an attempt to describe the quality of warehouse operation in terms of dependability. Authors discussed issues related to assessing warehouse operation, quality problems and solutions to increase the quality of work. The technical and organizational redundancy was proposed as a primary factor increasing dependability of warehous operation in supply chain and thereby improving the quality of services. Authors discussed dependability of supply chain and warehouses and have proposed an approach to determination of dependability of warehouse facility based on technological and organisational redundancy related to material flow pile-ups. The approach was founded on OTIFEF index as a base for dependability estimation. Construction of that index basing on probabilities of correct realization of different aspects of logistics service was proposed. An important element of the approach presented in the paper is proposal of technical and organisational indicators defining different aspects of redundancy in aspect of dependability. The example of redundancy assessment in function of technical and organisational methods of increasing warehouse efficiency has been provided.

Keywords: *warehouse dependability, technological redundancy, organizational redundancy, material flow pile-up, supply chain.*

Jakość pracy łańcucha dostaw jest wynikiem jakości pracy jego elementów, w tym magazynów. W artykule przedstawiono próbę ujęcia zagadnień jakościowych pracy magazynu w kategoriach niezawodnościowych. Omówiono zagadnienia oceny pracy magazynów, źródła problemów jakościowych i stosowane rozwiązania zwiększające jakość pracy. Wskazano nadmiarowość techniczną i organizacyjną jako podstawowy środek zwiększania niezawodności realizacji zadań przez magazyny w łańcuchu dostaw i tym samym poprawę jakości świadczonych usług. Autorzy omówili zagadnienia niezawodności łańcucha dostaw i magazynów jako ich podstawowych elementów i zaproponowali podejście do określania niezawodności magazynu oparte o nadmiarowość technologiczną i organizacyjną ustalaną w oparciu o przewidywane spiętrzenia przepływu materiałów. Do tego celu wykorzystano miernik OTIFEF jako podstawę szacowania niezawodności. Zaproponowano konstrukcję tego miernika w oparciu o prawdopodobieństwa poprawnej realizacji różnych aspektów usług logistycznych. Ważnym elementem podejścia proponowanego w artykule jest propozycja technicznych i organizacyjnych wskaźników określających różne aspekty nadmiarowości w funkcji niezawodności magazynu. Przedstawiono przykład szacowania nadmiarowości z wykorzystaniem technicznych i organizacyjnych metod zwiększania efektywności.

Słowa kluczowe: *niezawodność magazynu, nadmiarowość technologiczna, nadmiarowość organizacyjna, spiętrzenia w przepływie materiałów, łańcuch dostaw.*

1. Introduction

A key feature of any technical system, including logistics system, is work quality influencing user satisfaction. The quality of logistics services may be considered in relation to the time of delivery, security, dependability etc. in relation to costs [9, 11]). The quality of a logistic system services can be considered using dependability issues, but applying classic measures of dependability is impeded for logistics systems due to their complexity and necessary process-based approach to research.

Supply chain is a specific case of a logistics system. According to the serial structure of supply chain its quality, especially in technical matters, depends on quality of individual components [5]. These components – facilities and subsystems – perform processes of material transport, buffering and transformation.

Transport subsystems in supply chain determine efficiency and duration of material movement. They can be also a source of delays, damages and loss of materials. Warehouses buffer and transform materials as well as hold and deploy stock. Thus, warehouses determine accessibility of materials for clients and time of response for order. They are also places where the smallest possible pieces of materials in supply chain are touched (handled and transformed), so warehouses are potential sources of qualitative and quantitative errors. Warehouse processes are affected by the risk of damage to materials. Storage

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

process lowers value of materials and generates costs of maintaining inventory. These negative factors must be investigated and eliminated in order to improve the quality of services provided throughout the supply chain.

Therefore, warehousing seems to be important for the quality of services in the supply chain due to the broad spectrum of possible adverse events and errors caused by warehouses and identified by final customer. In case of transport subsystem shortcomings of performance can be quickly counteracted by using additional means of transport. In case of warehouses storage and handling capacity often cannot be achieved through simple organizational methods. In that case performance affects the timeliness and this, in turn, determines the quality.

Irregularities in logistics services result from errors of supply and distribution planning (not considered in this paper), technical deficiencies and inadequate organization. Hence we come up to the question as to the way that warehouses should operate, and in particular which technical means should they be fitted with to allow the undisturbed progress of warehousing processes in accordance with customer expectation? One of the elementary methods of enhancing dependability of systems is *redundancy*. Redundancy is understood as the application of a larger number of elements as compared to what is necessary or generally accepted. Those elements may comprise additional devices, people, space or information to be assured in case of damage or lack of efficiency of a system.

In the paper the authors have proposed a certain approach to warehouses services quality included in dependability categories. On that base they have discussed possibility of warehouse dependability improvement based on technological and organisational redundancy. In this context they have presented the research oriented at the dependability of logistic systems, particularly taking into account technical solutions in implementation of warehouse processes and supply chains. An important element of the presented approach is model presented in 5. point to determine selected characteristics of warehouse dependability in relation to material flow pile-ups. Additionally, authors propose factors of technical and organisational redundancy included into dependability structure of the warehouse.

2.Warehouse and supply chain dependability issues

In accordance with standards PN-82/N-04001 and PN-93/N-50191, dependability of a system (technical facility) is generally defined as a set of features that describe the readiness of the facility as well as reliability, maintainability and assuring maintenance support that affect it. According to Nowakowski [19, 20], when defining the dependability of a logistic system, the following should be taken into account:

- dependability of this system is understood only as a measure of task implementation over time, which may be compared to the reliability of technical system,
- no equivalent of maintainability or reliability of technical system has been formulated for logistics systems,
- all measures have the nature of coefficients, as a rule structure indicators; no other characteristics have been made use of, even though processes subject to assessment may also be random ones.

In other place Nowakowski [19] defines the dependability of logistics system through its availability. In a classic form function of availability of recoverable object describes the probability of its proper functioning in a specific moment of time [34]. Warehouse *availability* describes its ability of timely undertaking and successful performance of tasks arising from supply chain. It depend on the availability of resources, i.e. people, devices and means of internal transport, available time, control and measuring devices and buffering capability ([16], as per *Logistics Management Institute*). Dependability of the warehouse is a function of availability to handle supplies (unloading, receiving and put- away in the reserve area) as well as shipment (retrieving, replenishment, picking, sorting, consolidation, loading). The level of logistic services executed by the warehouse may be considered to be sufficient, and consequently the availability of the warehouse satisfactory, if at the fixed level of costs:

- number and structure of qualitative errors are at an acceptable level,
- services are provided on time, in accordance with the contract concluded with the customer,
- acceleration or cancellation of customer orders is possible in specific conditions,
- the facility is capable of handling pile-ups in material flows on time, without an adverse consequence for the remaining participants of the supply chain,
- the warehouse may, within preset limits, respond to changes in the goods structure or change in the structure in customer orders without increasing the number of occurring errors and maintaining the required timeliness.

Dependability is no easily measurable factor considered in designing logistic systems [10, 18]. In supply networks the issues of dependability are most frequently limited to the supply process. In this respect Nowakowski [19, 20] indicates that dependability in the operation of supply chains may indicate: timely task execution, complete implementation of an order and receipt or release of undamaged goods. A measure of functioning dependability of the system are disturbances and reduction of its performance [10, 18]. In this context the analysis comprises the flexibility of the system, i.e. the ability to adapt and to deploy changes in the operating scope and possibilities of increasing the operation area.

Numerous studies related to problems of dependability in supply chains [1-3, 8], take into account issues related to reducing system efficiency and changes in the loading of transport routes within the network [24, 27]. Sohn and Choi [28] analyse issues related to managing a supply chain in relation to the reliability of subsequent stages – logistic processes, including warehouse processes. They emphasise the need to include reliability issues already at the stage of designing. Bukowski and Feliks [6] search for a unified concept to evaluate dependability of complex supply chain. Baghalian et al. [1] present a mathematical model of the forming of broad assortment supply chains, which takes into account uncertainty on the demand and supply side. This uncertainty translates into material flow pile-ups on warehouse entrance and exit.

Quite interesting research of the impact exerted by the seasonality on the dependability level in transport services of spare parts was presented by Juściński and Piekarski [12]. The authors have set out the value of seasonal indices for the given period and analysed the distribution of their changes in the aspect of anticipated demand for spare parts for machines.

Issues of strategic management of the supply chain and planning of infallible warehouses functioning in regular conditions, as well as in conditions of unforeseen disturbances, have been subject of analyses performed by numerous authors [21, 25]. Peng et al. [21] presented a model for minimising the cost of logistic tasks implementation with concurrent minimising of the risk that the performance of those tasks would be discontinued. Neo et al. [15] analyse the impact of lack of warehouse technical efficiency on selected criteria of quality assessment of its operation. Furthermore, they point to the accuracy of information order-picking processes and timely execution of deliveries as crucial quality indicators.

Rizzi and Zamboni [23] regard logistic processes in manual warehouse of finished goods using ERP system to improve the productivity of the warehouse. The authors have pointed to the fact that the deployment of an integrated IT system is not guarantee for rationalising warehouse operation.

Werbińska-Wojciechowska [33] presents a model of maintaining technical systems on the example of logistic systems using the concept of time delays. Author points to the effectiveness of the devised model on the example of internal transport devices. In other work Werbińska-Wojciechowska [32] discusses the integration of the system executing the task with the supportive system like the logistics system.

Quality of services has considerable importance as it comes to dependability and is rather significantly highlighted by certain authors [26, 29, 30]. The authors perceive quality as a level at which customer requirements are met by services providers. For example in the study [26] as the Author describes the multidimensional nature of the quality, he states that the quality assessment should concern the dependability of operation of transport means and human resources, IT resources, infrastructure, technical equipping, and rules of system organisation. On the other hand, Chung and others [7] take up the issue of work quality and the dependability of warehouses in relation to the specialisation of the supply chain and the consequent possibility of re-using package resources. The authors propose the use of genetic algorithms as a tool for effective planning of dependability.

The crucial feature of dependability of supply chain and warehouse is the determination of the whole system faultless probability. This is difficult for complex structures like logistic systems, in the case of which classical damage causing lack of fitness of use is not applicable. Often warehouse processes are significant source of errors. In this context the appropriate number and selection of technical means has a particular significance. For this reason, it is necessary to specify selected technological and organisational measures for the selection of technical means to secure the implementation of warehouse processes. The selection must take into consideration the diversity of orders obtained from clients, and the irregularity of supplies.

On the basis of the performed review of literature an assumption may be made that there is a research gap on the impact of redundancy on dependability and the operating cost of the warehouse system, especially as regards technical solutions that may be allowed for use during designing of warehouse systems – when the precise progress of processes described by the history of order implementation and supply handling is not available.

3. Technological and organizational redundancy versus warehouse dependability

Review of literature and observations made in practice indicate that the *lack of dependability* of the warehouse process is connected with the following:

- human errors that generate repair tasks (additional labour, cost and delay), cause inconsistency of stock and reduce the quality of services in the supply chain,
- unforeseen pile-ups in the flow of materials causing lack of handling of current supplies and shipments,
- seasonal and long-term changes in the structure of supplies and shipments, which cause an insufficiency in performance and lack of technological suitability for the tasks,
- unreliability of mechanical equipment and facilities – especially the key elements of storage systems, such as AS/RS,
- warehouse and enterprise information systems errors (mismatching), poor quality planning.

Errors in warehouse process may be of two types: of *internal* or *external* nature in relation to the system. The first type of errors most frequently are an effect of human errors, an objective shortage of efficiency (productivity), mismatching of the technology and the tasks, shortage of buffering capacity and storage, and random events. External ones, from the viewpoint of the warehouse, arise exclusively from random causes connected with changes in changes to the supply structure. The nature and scope of those changes is not reviewed at the warehouse level but is subject of planning of the supply chain (Fig. 1).

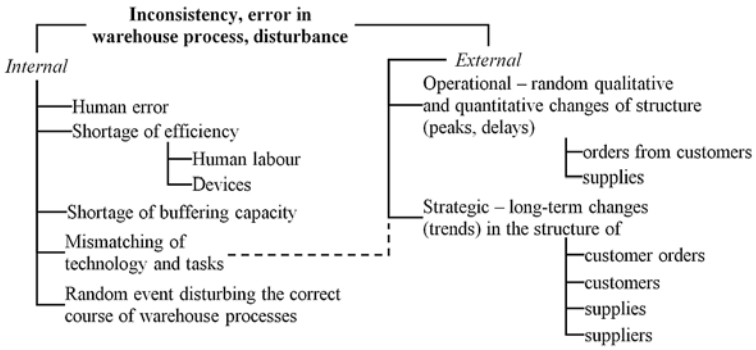


Fig. 1. Classification of sources of errors and hindrances that limit the warehouse dependability. Source: [14]

The conducted research [4, 16, 17] clearly points to a necessity of carrying out an analysis of a sample of historical data with view to dependability to set out measurable indicators related to the number and types of possible errors. On that base warehouse technologies and methods of work organisation are selected to reduce the number or change the structure of errors in the implementation of warehouse processes.

Problems in implementation of warehouse processes, such as the lack of punctuality, may be a result of the absence of handling potential or lack of suitability of the potential to random pile-ups in material flows [31]. A negative influence of random peaks in daily material flows on its uninterrupted operation is flattened by daily pile-up coefficients expressing technological redundancy allowed for in the designing stage. The coefficients increase average daily volumes to set reliable (nominal) material volumes covering the majority of daily flow peaks (section 5). The application of pile-up coefficients that describe processes of supplies and shipments enables taking into account the efficiency overcapacity in planning of warehouse processes, and consequently enhancing the dependability of the facility.

The pile-up coefficients do not take into account extraordinary situations or long-term trends in the supply chain that arise from the seasonal change in the structure of flows and supplies, and long-term changes on the market. The basic way of coping with them is the periodical change in work organisation and adaptation of efficiency. Long-term changes tend to reduce the availability to providing services and require the adoption of flexible solutions offering the given possibility of adaptation which, in turn, are connected with the technological universality and with the anticipated redundancy (installed in practice and potential one).

Typical methods of eliminating the above mentioned issues are as follows:

- *technological redundancy* enhancing the dependability of the system, including
  - increasing the capacity of functional areas (specially storage),
  - increasing number of people and equipment,
  - using equipment more efficient than actually required.
- *technological universalism* (*flexibility*),
- *organisational changes* that extend the available time of work and which allow a reduction of the required volume of resources,
- *adoption of integrated management systems* (WMS) to eliminate out-of-system activities,
- *adoption of motivation programmes* of the employees.

A review of the following aspects allows the presumption that planning of a warehouse facility requires taking into account redundancy of the following nature:

- *Functional* (flexibility, universalism), i.e. possibilities of functional reconfiguring of the system to allow its adaptation to the nature of implemented logistic tasks.
- *Technological* (oversizing), i.e. an overcapacity in efficiency as compared to needs.

Redundancy always has to be confronted with the effectiveness of the warehouse and whole supply chain. Oversizing to exceed actual needs causes a high unit cost of tasks, while universalism may cause a drop in competitiveness. The adequacy of that cost arises from market conditions.

Organisational methods aimed at increasing the dependability of a warehouse are based on the following:

- rational allotment of resources for needs of warehouse tasks in time,
- determination of available time for task implementation during the day,
- selection of universal devices which may be moved between tasks,
- adoption of methods directing the flow of materials to and from the given location to reduce the work intensity of the process, in the function of costs of task implementation and availability of resources.

The dependability of the warehouse may be periodically enhanced by the adoption of organisational means aimed at increasing the utilisation rate of the working time (e.g. by motivating) or permanently by application of warehouse management systems. Temporary factors that enhance efficiency, and consequently improve the overall dependability of the warehouse in the supply chain include the possibility of extending the daily working time. Long-term increase in performance by extending the daily working time requires employing additional people.

Functional redundancy is also contained in schedule of warehouse process that describes work pile-ups. Planning the process in a way that increases the available time of task implementation concurrently maintaining the same state of resources would allow enhancing the system potential [13].

Both technological and functional redundancy are indispensable for proper realization of warehouse processes. Certain aspects of technological and functional redundancy may be taken into consideration in the project phase, as was outlined below.

4. Synthetic warehouse dependability measure

Warehouse dependability can be related to the basic features of properly realized logistics service defined by 7R rule [19], which means *transforming entering material flows (from suppliers) into materials for clients according to orders, within agreed time, with adequate quality and costs.* Quality of warehouse services is sufficient, and consequently its dependability is satisfactory, when number and structure of qualitative errors is admissible by client, services are provided on time and warehouse is capable of handling predictable and non-predictable pile-ups in material flows. Then, it is possible to determine overall dependability measure *OTIFEF* (*On-Time, In-Full, Error-Free*) of warehouse on the base of warehousing operations history [14, 19]. *OTIFEF* measure is basically the ratio of tasks (services) completed in a way fully satisfying customers (model) to the number of all ordered tasks.

Boundary conditions of warehouse operation are determined by structure of deliveries on one side and structure of shipments resulting from customer orders on the other. Customer satisfaction must go hand in hand with correct servicing of suppliers. Thus, $OTIFEF_{in}$ describes quality of servicing suppliers (feeding warehouse) and $OTIFEF_{out}$ describes quality of shipments (emptying warehouse) [14]. Both measures are defined by a set of parameters determining quality of work, but generally can be composed as follows:

$$\forall m \in \boldsymbol{M} \quad OTIFEF_{in}(m) = P_{OTin}(m) \cdot P_{IFin}(m) \cdot P_{EFin}(m) \tag{1}$$

where:

$P_{OTin}(m)$ – probability of handling all (daily) supplies on-time in m-th warehouse,
$P_{IFin}(m)$ – probability of handling all (daily) supplies in-full in m-th warehouse,
$P_{EFin}(m)$ – probability of handling all (daily) supplies with no errors in m-th warehouse.

A similar function is formulated for shipments:

$$\forall m \in \boldsymbol{M} \quad OTIFEF_{out}(m) = P_{OTout}(m) \cdot P_{IFout}(m) \cdot P_{EFout}(m) \tag{2}$$

where:

$P_{OTout}(m)$ – probability of handling all (daily) shipments on-time in m-th warehouse,
$P_{IFout}(m)$ – probability of handling all (daily) shipments in-full in m-th warehouse,
$P_{EFout}(m)$ – probability of handling all (daily) shipments with no errors in m-th warehouse.

For the purpose of research, probabilities are considered as independent. This can result in underestimation of warehouse dependability expressed by formulas (1) or (2), but is acceptable when functions are used for comparing technical and organizational variants.

Probabilities of handling all shipments and supplies on-time P_{OT} are directly dependent on technical potential od warehouse. Probabilities of quality error P_{EF} are related to human factor and random events. Probabilities P_{IF} of servicing daily supplies and shipments in-full are related to the availability of free storage place and ordered materials on hand, which are dependent on inventory planning strategies and indirectly on storage capacity.

5. Reliable material flow volumes vs punctuality

The fundamental step of designing warehouse is determining reliable (nominal) material flow volumes on entrance (supplies) and on exit (shipments). Reliable material flows constitute the base for counting number of handling equipment, workers and spaces. Installed technical potential must be able to handle all daily supplies and shipments and random pile-ups in material flow volumes.

Volumes of materials entering the warehouse and leaving it can be described by the relevant distribution (example for supplies in Figure 2).

Daily material flow volumes on entrance are described by random variable $\{\lambda_{in}(m),\ p(\lambda_{in}(m))\}$ and on exit (shipments) by $\{\lambda_{out}(m), p(\lambda_{out}(m))\}$. Flow volumes are expressed by number of handled unified units. Average daily flow volumes are then set as expected values:

$$\forall m \in \boldsymbol{M} \quad \lambda_{in}^{av}(m) = E(\lambda_{in}(m)) \text{ and} \tag{3}$$

$$\forall m \in \boldsymbol{M} \quad \lambda_{out}^{av}(m) = E(\lambda_{out}(m)) \tag{4}$$

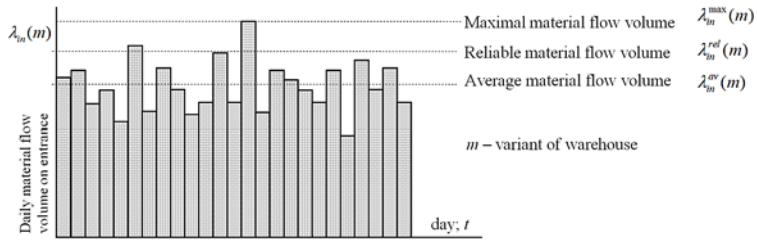


Fig. 2. Distribution of material flow volume on warehouse entrance (supplies). Source: own work

Reliable flow volumes are the expected volumes that must be serviced on-time, even under statistically expected equipment breakdowns and workers' unavailability (necessary redundancy). Finding reliable flow volumes is difficult and depends on type of warehouse, supply chain organization and serviced business.

Random variables $\{\lambda_{in}(m), p(\lambda_{in}(m))\}$ and $\{\lambda_{out}(m), p(\lambda_{out}(m))\}$ are characterized by coefficients of variation V:

$$V_{in}(m) = \frac{\sigma_{in}(m)}{\mu_{in}(m)} \text{ and } V_{out}(m) = \frac{\sigma_{out}(m)}{\mu_{out}(m)} \tag{5}$$

where μ is the mean and σ is the standard deviation.

Coefficients of variation V are different for different material flow strategies in supply chain, but observations and analyses revealed reference values (Table 1).

Table 1. Exemplary variation coefficients of material flow volumes in warehouse facility.

Type of warehouse	Minimal V	...	Maximal V
Industry (production) warehouse entries (supplies)	0,03	0,07	0,12
Industry (production) warehouse exit (shipments)	0,02	0,05	0,1
Distribution (retail) warehouse entries (supplies)	0,2	0,35	0,65
Distribution (retail) warehouse exit (shipments)	0,35	0,45	1,12

Source: own research.

The variation coefficient $V < 0{,}1$ is considered as irrelevant, which is reflected in the way of setting reliable material flow volumes on entrance and on exit:

$$\forall m \in \boldsymbol{M} \text{ if } \begin{cases} V_{in}(m) \le k1_{in}(m) \rightarrow & \lambda_{in}^{rel}(m) = \lambda_{in}^{\max} = \max\{\lambda_{in}(m)\} \\ V_{in}(m) > k1_{in}(m) \rightarrow & \lambda_{in}^{rel}(m) = E(\lambda_{in}(m)) + k2_{in}\sqrt{E((\lambda_{in}(m))^2) - E(\lambda_{in}(m))^2} \end{cases} \tag{6}$$

and

$$\forall m \in \boldsymbol{M} \text{ if } \begin{cases} V_{out}(m) \le k1_{out}(m) \rightarrow & \lambda_{out}^{rel}(m) = \lambda_{out}^{\max} = \max\{\lambda_{out}(m)\} \\ V_{out}(m) > k1_{out}(m) \rightarrow & \lambda_{out}^{rel}(m) = E(\lambda_{out}(m)) + k2_{out}\sqrt{E((\lambda_{out}(m))^2) - E(\lambda_{out}(m))^2} \end{cases} \tag{7}$$

where:

$k1_{in/out}(m)$– variation coefficient determining irrelevancy for particular decision situation,

$k2_{in/out}(m)$– index representing required warehouse service level.

Low V means that material flow is not disturbed by random pile-ups and is non-changeable so reliable flow volume is equal to maximal flow volume. This situation is characteristic for warehouses handling high volumes of unified, non-seasonal products (like feeding production or cross-docking). High variation coefficient V means that warehouse experiences high pile-ups in material flows. In many cases pile-ups appear rarely and don't justify installing redundant handling potential. These pile-ups excessing reliable flow volumes must be serviced by extra potential of three types, which may be combined (Table 2).

Table 2. The possibilities of serving material flow pile-ups exceeding reliable material flows.

Event	Description
A	universal equipment and workers can be moved from other tasks and places in warehouse to handle high pile-ups on entrance or on exit
B	daily work time can be extended (additional FTE) or short-term improvement of work time utilization through motivation methods can be used
C	extra equipment, storage space and human resources can be gained from outside (renting equipment and space, hiring temporary workers)

Source: own research.

Hence values of key importance for warehouse designing are the pile-up coefficients:

$$\forall m \in \boldsymbol{M} \quad \varphi_{in}(m) = \frac{\lambda_{in}^{rel}(m)}{\lambda_{in}^{av}(m)} \text{ and} \tag{8}$$

$$\forall m \in \boldsymbol{M} \quad \varphi_{out}(m) = \frac{\lambda_{out}^{rel}(m)}{\lambda_{out}^{av}(m)} \tag{9}$$

Therefore, warehouse facility can be described by important dependability characteristics:

$$\forall m \in \boldsymbol{M} \quad P\left(\frac{\lambda_{in}(m)}{\lambda_{in}^{av}(m)} \le \varphi_{in}(m)\right) = \alpha_{in}(m) \text{ and} \tag{10}$$

$$\forall m \in \boldsymbol{M} \quad P\left(\frac{\lambda_{out}(m)}{\lambda_{out}^{av}(m)} \le \varphi_{out}(m)\right) = \alpha_{out}(m) \tag{11}$$

It was assumed that $\alpha_{in}(m) = \alpha_{out}(m) \ge 0{,}96$ is a typical value, but it depends on the analysed case. Number of workers and equipment and storage capacities are set to meet determined material flow volumes. Assuming certain simplifications and invariability of technology, it can be stated that these numbers are proportional to the daily flow volume.

Concluding formulas (1), (2) and (6), (7) warehouse dependability in terms of punctuality can be considered in aspect of technical redundancy in two ways:

- event H_{1in} when $\lambda_{in}(m) \le \lambda_{in}^{rel}(m)$ – flow volumes on entrance are below the limit, so can be handled on-time in 100%.
- event H_{2in} when $\lambda_{in}(m) > \lambda_{in}^{rel}(m)$ – flow volumes on entrance are over the limit, so can't be handled on-time in 100%.
- event H_{1out} when $\lambda_{out}(m) \le \lambda_{out}^{rel}(m)$ – flow volumes on exit are below the limit, so can be handled on-time in 100%.

Table 3. Structure of errors in order-picking process in company X.

Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct
No of picked lines	31 360	18 722	12 095	16 802	11 991	9 988	11 489	7 119	7 939	11 425
Number of errors	54	26	48	43	23	47	13	38	26	30
- quality errors	28	16	32	29	16	32	11	25	15	20
- quantitative errors	26	10	16	14	7	15	2	13	11	10

Source: own research.

– event H_{2out} when $\lambda_{out}(m) > \lambda_{out}^{rel}(m)$ – flow volumes on exit are over the limit, so can't be handled on-time in 100%.

If $\lambda_{in}^{rel}(m) = \lambda_{in}^{\max}(m)$ for m-th warehouse (formula (6)), then:

$$P_{OTin}(m) = P(H_{1in}) \approx 1, \tag{12}$$

and similarly if $\lambda_{out}^{rel}(m) = \lambda_{out}^{\max}(m)$ for m-th warehouse (formula (7)), then:

$$P_{OTout}(m) = P(H_{1out}) \approx 1. \tag{13}$$

If $\lambda_{in}^{rel}(m) < \lambda_{in}^{\max}(m)$ for m-th warehouse (formula (6)), then: ways of coping with pile-ups presented in Table 2 are applied:

$$\begin{aligned} P_{OTin} &= P(H_{2in}/A)\cdot P(A) + P(H_{2in}/B)\cdot P(B) + P(H_{2in}/C)\cdot P(C) + \\ &+ P(H_{2in}/A\cap B\cap C)\cdot P(A\cap B\cap C) + P(H_{2in}/A\cap B)\cdot P(A\cap B) + \\ &+ P(H_{2in}/A\cap C)\cdot P(A\cap C) + P(H_{2in}/B\cap C)\cdot P(B\cap C) \end{aligned} \tag{14}$$

where $P(A)$, $P(B)$ and $P(C)$ are appropriate probabilities describing availability of particular solutions in analyzed warehouse at the time of pile-ups.

Similarly, if $\lambda_{out}^{rel}(m) < \lambda_{out}^{\max}(m)$ for m-th variant of the warehouse (formula (7)), then:

$$\begin{aligned} P_{OTout} &= P(H_{2out}/A)\cdot P(A) + P(H_{2out}/B)\cdot P(B) + P(H_{2out}/C)\cdot P(C) + \\ &+ P(H_{2out}/A\cap B\wedge C)\cdot P(A\cap B\cap C) + P(H_{2out}/A\cap B)\cdot P(A\wedge B) + \\ &+ P(H_{2out}/A\cap C)\cdot P(A\cap C) + P(H_{2out}/B\cap C)\cdot P(B\cap C) \end{aligned} \tag{15}$$

The high capacity of pile-ups handling requires maintaining a considerable technological potential in readiness – redundancy. The possibility of handling smaller pile-ups generate lower costs, yet it would also limit the readiness of the warehouse to provide services. Technological redundancy resulting from handleable pile-ups in material flow is reflected in organizational and technical indicators of warehouse system.

6. Structure of qualitative errors and availability of materials

As described in section 4. synthetic measure of warehouse dependability *OTIFEF* is based on three components. The first component (*On-time*) is discussed in section 5. The other two are difficult to be measured and require data mining.

Structure of qualitative errors made by the employees (component *Error-free*) can be estimated only through analysis of historical data of warehouse operation or using methods like FMECA (see [33]). Analysis of historical data allows to determine the probability distributions of qualitative errors under specific technological and organizational configuration of warehouse and known workload. Of course, there are technologies and solutions that contribute to minimizing the number of errors – especially automated technologies and strict control of the process by the warehouse information system like WMS. Unfortunately, the effect of the implementation of such technologies can be determined only on the basis of analysis of their work.

An analysis of the structure of errors in order picking process in real warehouse was performed to illustrate the problem. The analysis covered a period of 10 months, during which 138 930 lines of orders was realized (Table 3).

The analysis has revealed that probability distribution of correct execution of the orders can be approximated by Weibull distribution. Estimated expected probability of error-free implementation of single order-line in terms of quality is 0.9972656 and in terms of quantity is 0.9984898. The probability of faultless execution of order-line is then 0.9957417. The probabilities set in that way can be used to determine the $OTIFEF_{out}$ index. It must be noted that these data do not include errors that were identified before sending materials to customers.

The last component of the measure – *In-full* is a function of the availability of space for storage units incoming to the warehouse and on-hand availability of ordered materials to be released. It results from the supply and distribution planning mechanisms, product features and warehouse storage area capacity. In addition, the products in most types of businesses can be divided into fast and slow-moving. In typical situations, fast moving materials are likely available immediately (make-to-stock), while the slow-moving products may not be available at the time of placing an order (make-to order). It is thus possible to determine the probabilities of execution of orders "in-full".

7. Selected technical and organizational indicators of warehouse system redundancy

7.1. Elements of dependability structure of warehousing system

Basic criteria for assessing the quality of a warehouse as an element of the supply chain should comprise technical measures (productivity, performance), economic measures (costs and investment expenditures) and qualitative measures (number and structure of errors). These measures are useful for determining dependability in terms of redundancy. All elements of the supply chain must meet separate expectations as to efficiency. This indicates that elements within the chain may be perceived as the dependability system of the entire series. Lack of reliability of one or more elements is transposed on the lack of reliability of the entire supply chain. And vice versa, dependability of particular elements of the supply chain causes that it may meet expectations related to efficiency on all markets in its surroundings.

Hence, if we assume that the analysed supply chain is of a serial structure and we know:

– set of elements of supply chain $\boldsymbol{V} = \{v : v = \overline{1,V}\}$,

– set of connections between elements $\boldsymbol{L} = \{(v,v') : v,v' \in \boldsymbol{V} \wedge (v,v') \in \boldsymbol{V} \times \boldsymbol{V} \wedge v \neq v'\}$
– dependability indicator of the v-th elements of supply chain: $nl(v)$;
– dependability indicator of connection between links (v, v') of supply chain: $nl(v, v')$,

dependability in the structural sense for the entire chain may be determined as following:

$$WNS = \prod_{(v,v')\in \boldsymbol{L}} nl(v,v') \cdot \prod_{v\in \boldsymbol{V}} nv(v) \tag{16}$$

Furthermore, taking into account the routes implemented in supply chains, assuming that the following data have been determined:
– set of sources of material flow in supply chain: $\boldsymbol{A}$,
– set of destinations of material flow in supply chain: $\boldsymbol{B}$,
– set of numbers of routes which may join sources with end nodes: $\boldsymbol{E}(a,b)$, $a \in \boldsymbol{A}$, $b \in \boldsymbol{B}$
– set of numbers of nodes for all routes: $\boldsymbol{EW}(a,b,e)$, $a \in \boldsymbol{A}$, $b \in \boldsymbol{B}$, $e \in \boldsymbol{E}(a,b)$
– set of edges determining e-th flow itinerary in supply chain in relation (a, b): $\boldsymbol{EL}(a,b,e)$, $a \in \boldsymbol{A}$, $b \in \boldsymbol{B}$, $e \in \boldsymbol{E}(a,b)$.

dependability of the supply chain in the sense of transport and warehouse routes may be defined in the following way:

$$WNC = \prod_{a\in \boldsymbol{A}} \prod_{b\in \boldsymbol{B}} \left[\sum_{e\in \boldsymbol{E}(ld,a,b)} \left(\prod_{(v,v')\in \boldsymbol{EL}(a,b,e)} nl(v,v') \cdot \prod_{v\in \boldsymbol{EW}(a,b,e)} nv(v) \right) \right] \tag{17}$$

Ascribing of highly reliable resources or excessive number of resources to the realization of particular routes in supply chain increases the operability of the entire chain.

7.2. Selected technical and organisational measures of redundancy of warehouse system in aspect of task implementation dependability

It was assumed that given supply chain uses M warehouses. Set of numbers of warehouses is denoted as $\boldsymbol{M} = \{m : m = \overline{1,M}\}$. Each warehouse uses a set of resources (equipment and workers). A single type of resource in m-th warehouse is marked as r, so the set $\boldsymbol{R}(m)$ of resources is described as following:

$$\forall m \in \boldsymbol{M} \quad \boldsymbol{R}(m) = \{r:\ r = 1, 2, \ldots, R(m)\}.$$

It was assumed that the warehouse process consists of sequentially numbered transformations of material flows $i, j \in \boldsymbol{I}(m)$ implemented by resources $r \in \boldsymbol{R}(m)$. If $\alpha(r, m, i) = 1$ then r-th type of resource implements i-th task in m-th warehouse. The daily work-load of the r-th type resource is the sum of products of the number of transport operations under i-th tasks $\lambda(m, i)$ and the duration of a single reiteration $t(r, m, i)$ by r-th resource:

$$\forall m \in \boldsymbol{M}\ \forall r \in \boldsymbol{R}(m) \quad \Psi(r,m) = \sum_{i\in \boldsymbol{I}(m)} \alpha(r,m,i) \cdot \lambda(m,i) \cdot t(r,m,i) \tag{18}$$

where: $\forall m \in \boldsymbol{M}\ \forall i \in \boldsymbol{I}(m)\ \ \lambda(m, i) = \lambda^{av}(m, i) \cdot \varphi_{in}(m)$ or $\lambda(m, i) = \lambda^{av}(m, i) \cdot \varphi_{out}(m)$ (19)

depending on whether i-th task is for handling supplies (in) or shipments (out).

Each type of resource has a specified cost of hourly operation $k(r,m)$. Balancing the workload $\Psi(r,m)$ with the cost of operation allows obtaining standardised workload, which may be added for all types of equipment:

$$\forall m \in \boldsymbol{M} \quad \widetilde{\Psi}(m) = \sum_{u\in \boldsymbol{R}(m)} \frac{\Psi(r,m) \cdot k(r,m)}{\min\{k(r,m) : r \in \boldsymbol{R}(m)\}} \tag{20}$$

The number $n(r, m)$ of r-th resources is known. Standardised number of resources may be determined, which is to define the technological redundancy of the warehouse as compared to other variants of the system:

$$\forall m \in \boldsymbol{M} \quad \bar{n}(m) = \sum_{r\in \boldsymbol{R}(m)} \frac{\alpha(r,m,i) n(r,m) k(r,m)}{\min\{k(r,m) : r \in \boldsymbol{R}(m)\}} \tag{21}$$

Each task has an assigned available implementation time $td(i, m)$ arising from the daily warehouse operation schedule. The tasks may be implemented concurrently, which as a result leads to overlapping of tasks in certain periods and work piling up.

Indicator of available operating time utilization

Consequently, the technological redundancy of a warehouse system that determines its dependability may be expressed by the indicator of available operating time utilization of technical resources of the r-th type:

$$\forall m \in \boldsymbol{M} \quad \theta(r,m) = \frac{\sum_{i\in \boldsymbol{I}(m)} \Psi(r,m)}{n(r,m)\varphi(r,m) t_{dob}(m)} \tag{22}$$

where:
$t_{dob}(m)$ – daily operating time of the m-th warehouse,
$\varphi(r, m)$ – operating time utilisation by r-th resources in m-th warehouse.

This indicator is a quotient of the work-load and daily operating time appointed to the resources of the given type. Controlling the daily operating time for task implementation is a basic organisational tool oriented at increasing the efficiency of the warehouse and used to handle non-standard pile-ups in material flow.

Cost-related organisational index

The potential technical redundancy of the warehouse system may also be expressed in cost categories, among others by the cost-based organisational indicator for assessing the utilisation level of installed devices:

$$\forall m \in \boldsymbol{M} \quad \theta^{K}_{oz}(m) = \frac{K^{Ro}(m)}{K^{R}_{T}(m) + K^{R}_{S}(m) + K^{R}_{L}(m)} \tag{23}$$

where:
$K^{Ro}(m)$ – annual operating costs (direct cost of labour and equipment usage) [PLN/year],
$K^{R}_{T}(m)$ – total annual maintenance costs (all costs, including depreciation) [PLN/year],
$K^{R}_{S}(m)$ – total annual maintenance costs of control systems [PLN/year],
$K^{R}_{L}(m)$ – total annual labour costs [PLN/year].

The optimum value of this indicator – 1, means that all resources are used in 100% during the working day. Lower values point to the

existence of a technological potential, which may be initiated by taking up appropriate organisational means. In most cases this potential is not used due to the defined allotment of resources and work organisation contributing to work pile-ups that involve all resources, which are not used in the remaining time.

Organisational index expressed by standardised work intensity

The organisational index expressed by standardised work intensity of process implementation may be described in the following way:

$$\forall m \in \boldsymbol{M} \;\; \theta_{oz}^{\widetilde{\Psi}}(m) = \sum_{r \in \boldsymbol{R}(m)} \frac{\Psi(r,m) \cdot k(r,m)}{\sum\limits_{r' \in \boldsymbol{R}(m)} k(r',m) t_{dob}(m) \max\limits_{\tau \in \boldsymbol{T}} \{\Psi(r,m,\tau)\}} \quad (24)$$

where: $\max\limits_{\tau \in \boldsymbol{T}} \{\Psi(r,m,\tau)\}$ is maximal temporary work intensity for the r-th type of resource arising from work pile-ups at τ-th moment during a 24-hour period [m.h/h],

The maximum temporary work intensity of the process $\max\limits_{\tau \in \boldsymbol{T}} \{\Psi(r,m,\tau)\}$ with view to operation of r-type resources is understood as a maximum sum of work intensity of successive i-th tasks implemented in a parallel way falling for the τ-th time interval. The distribution of task implementation during each 24-hour period arises from the schedule (organisation).

8. Example of determining selected redundancy measures for dependability assessment

The research was carried out for a warehouse executing processes composed of 14 tasks. Potential pile-ups in deliveries and shipments are set by the pile-up coefficients. The installed potential was analysed with view of effectiveness and efficiency for pile-up coefficients on entry φ_{in} and at exit φ_{out} as 1.1; 1.3 and 1.5 respectively. The warehouse has at its disposal resources presented in Table 4.

The warehouse is working one shift 290 days a year. There are no seasonal changes in material flows. The average size of daily reloading operations on entry: 300 pallet units. The average number of release operations: 455 (including consolidated and homogenous units). The workload of successive tasks of the process arise from the technology and geometry of the building. The floor area and the storage capacity remain unchanged. An analysis was performed of the performance and warehouse costs for he defined schedule and without it (allotting the entire daily working time for tasks) to determine the potential organisational reserve.

Results of efficiency of scheduled warehouse operation are presented in Table 5. Standardised work intensity of the process which

Table 4. Listing of types of devices (u) and categories of human labour (c) in the warehouse.

Type	Description	Q-ty	Cost of an hour of work [PLN/h] – net as regards employees	Utilisation degree of working time
$u1$	Powered lifting pallet truck	3	4.00	0.8
$u2$	Front lifting pallet truck	14	7.00	0.9
$u3$	Horizontal order picking trolley	9	9.00	0.9
$u4$	High reach truck	8	12.00	0.9
$c1$	Operator of $u1$, $u2$ + manual work	var.	13.00	0.8
$c2$	Employee for picking and control		16.00	0.8
$c3$	Operator of $u2$ and $u3$		20.00	0.8

Table 5. Technical parameters of warehouse operation – with schedule.

Parameter	Variant								
	1	2	3	4	5	6	7	8	9
φ_{in}	1.5	1.5	1.5	1.3	1.3	1.3	1.1	1.1	1.1
φ_{out}	1.5	1.3	1.1	1.5	1.3	1.1	1.5	1.3	1.1
Maximum intensity of standardised work intensity									
with view to operation of devices [m.h]	57.50	52.48	47.46	54.89	49.87	44.84	52.23	47.20	42.18
with view to labour of employees [m.h]	41.95	38.32	34.69	40.01	36.39	32.76	38.03	34.40	30.78
Standardised work intensity									
with view to operation of devices [m.h]	324.88	296.67	268.46	309.94	281.73	253.52	294.74	266.53	238.32
with view to labour of employees [m.h]	233.55	213.38	193.21	222.72	202.55	182.38	211.67	191.50	171.33
Organisational indicator #1 – effectiveness of utilisation of installed potential									
with view to operation of devices:	0.706	0.707	0.707	0.706	0.706	0.707	0.705	0.706	0.706
with view to labour of employees:	0.696	0.696	0.696	0.696	0.696	0.696	0.696	0.696	0.696
Organisational indicator #2 – unutilised technological potential (standardised work intensity)									
with view to operation of devices [m.h]	135.15	163.36	191.57	150.09	178.30	206.52	165.29	193.50	221.71
with view to labour of employees [m.h]	102.07	122.24	142.41	112.90	133.07	153.24	123.95	144.12	164.29
Organisational indicator #2 – unutilised technological potential (standardised work intensity)									
with view to operation of devices	29.4%	35.5%	41.6%	32.6%	38.8%	44.9%	35.9%	42.1%	48.2%
with view to labour of employees	30.4%	36.4%	42.4%	33.6%	39.6%	45.7%	36.9%	42.9%	49.0%

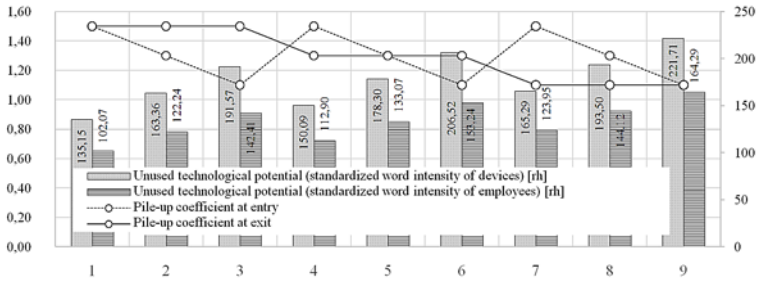


Fig. 3. Unused technological potential – with schedule

expresses the general volume of work to be executed in relation to work costs was used for assessment of the technological overcapacity of the warehouse. Unused technological potential in variant 1 (the highest pile-ups at level α which may be handled) arises from the structure of the work schedule, in which tasks may only be implemented in certain sections of the daily operation time.

The level of unused technical and functional potential (redundancy) was presented on Fig. 3.

Table 6 presents performance results of warehouse operation without schedule (all tasks are executed evenly throughout the entire daily working time – no pile-ups occur, resources are used in 100%). Actual and standardised work intensity are the same as if a schedule is applied.

Fig. 4. presents the technological redundancy for the process without the schedule. In variant 1 standardised work intensity is distributed evenly within the entire daily working time, hence the lack (0%) of technological and organisational overcapacity. In subsequent variants the technological redundancy arises from redundant resources, and no functional redundancy occurs.

Table 7 and Fig. 5. list costs of warehouses processes. Annual operating costs take into account all costs connected with the execution of warehouse processes and maintaining warehouse infrastructure. All other dependability and technical indicators have to refer to operating costs as the ultimate profitability index.

Fig. 6. presents costs of implementation of a single customer order depending on values of material pile-ups that may be handled. Also presented is the value of organisational indicator.

In case of lack of schedule and therefore daily pile-ups, investigated process may be carried out by the same resources (equipment) under pile-up coefficients on entry φ_{in} = 2,98 and at exit φ_{out} = 1,1 and in the inverse situation under φ_{in} = 1,1 and φ_{out} = 2,32. This reserve (initial values of pile-up coefficients $\varphi_{in} = \varphi_{out}$ = 1,1) is significant. Naturally it cannot be fully used due to the technological limitations of the process, which will force the schedule, however, it indicates a potential functional redundancy possible to run under certain conditions.

Table 6. Technical parameters of warehouse operation – without schedule.

Parameter	Variant								
	1	2	3	4	5	6	7	8	9
φ_{in}	1.5	1.5	1.5	1.3	1.3	1.3	1.1	1.1	1.1
φ_{out}	1.5	1.3	1.1	1.5	1.3	1.1	1.5	1.3	1.1
Maximum intensity of standardised work intensity									
with view to operation of devices [m.h]	40.61	37.08	33.56	38.74	35.22	31.69	36.84	33.32	29.79
with view to human labour [m.h]	29.19	26.67	24.15	27.84	25.32	22.80	26.46	23.94	21.42
Organisational indicator #2 – unused technological potential expressed by standardised work intensity in relation to the baseline solution (variant 1)									
with view to operation of devices [m.h]	0.0%	8.7%	17.4%	4.6%	13.3%	22.0%	9.3%	18.0%	26.6%
with view to human labour [m.h]	0.0%	8.6%	17.3%	4.6%	13.3%	21.9%	9.4%	18.0%	26.6%

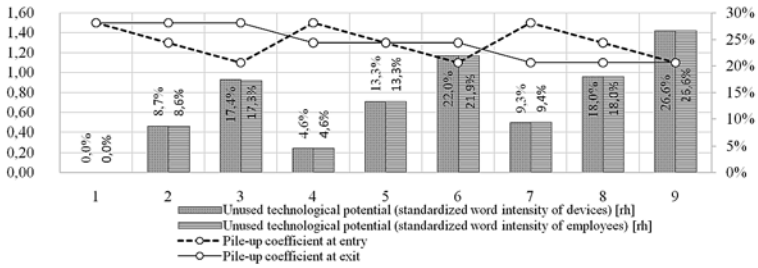


Fig. 4. Unused technological potential – with no schedule.

Table 7. Listing of cost parameters for assessment of warehouse effectiveness – with schedule.

Parameter	Variant								
	1	2	3	4	5	6	7	8	9
φ_{in}	1.5	1.5	1.5	1.3	1.3	1.3	1.1	1.1	1.1
φ_{out}	1.5	1.3	1.1	1.5	1.3	1.1	1.5	1.3	1.1
Annual operational cost of work									
of people [PLN million/year]	2.780	2.476	2.300	2.589	2.425	2.171	2.538	2.234	2.058
of equipment [PLN million/year]	0.420	0.384	0.347	0.401	0.365	0.328	0.382	0.345	0.308
Annual operating costs									
of equipment [PLN million/year]	4.928	4.891	4.854	4.908	4.872	4.835	4.889	4.852	4.816
of warehouse (total) [PLN million/year]	8.360	8.019	7.806	8.149	7.949	7.658	8.079	7.738	7.525
Cost-related organisational indicator	0.44	0.41	0.39	0.42	0.40	0.37	0.41	0.38	0.36
Effectiveness of solution [PLN/released unit]	42.33	46.31	52.47	41.63	46.37	52.08	41.70	45.68	51.90

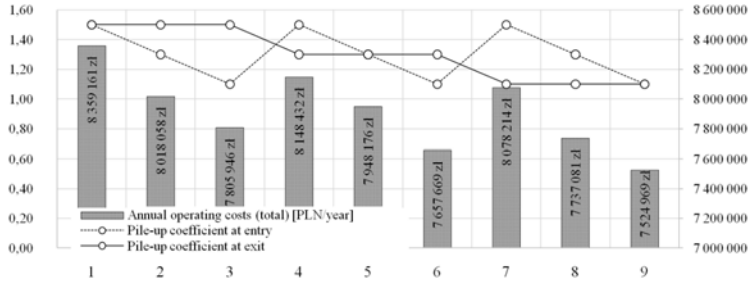


Fig. 5. Value of annual working cost for successive pile-up coefficients

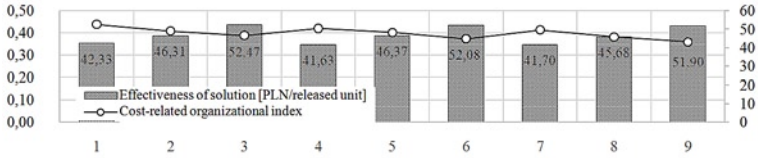


Fig. 6. Effectiveness of solutions in successive variants

Table 8. Technological redundancy determined by available time of equipment – with schedule.

Parameter	Variant								
	1	2	3	4	5	6	7	8	9
φ^{in}	1.5	1.5	1.5	1.3	1.3	1.3	1.1	1.1	1.1
φ^{out}	1.5	1.3	1.1	1.5	1.3	1.1	1.5	1.3	1.1
Total number of employees	46	41	38	43	40	36	42	37	34
Computational number of equipment elements [pieces]									
*u*1	2.27	2.00	1.73	2.24	1.97	1.70	2.21	1.94	1.67
*u*2	13.06	12.09	11.11	12.29	11.32	10.35	11.51	10.54	9.57
u3	8.93	7.86	6.79	8.82	7.75	6.68	8.70	7.63	6.56
u4	7.58	7.08	6.59	7.06	6.56	6.07	6.54	6.05	5.55
Computational number of employees [persons]									
*c*1	19.33	17.68	16.03	18.40	16.76	15.11	17.46	15.81	14.17
*c*2	16.25	14.60	12.94	15.76	14.11	12.45	15.24	13.59	11.94
c3	8.52	7.97	7.41	7.94	7.39	6.83	7.36	6.80	6.25
Surplus of performance expressed by standardised number									
with respect to equipment:	0%	7%	17%	3%	14%	19%	9%	16%	26%
with respect to employees:	0%	11%	17%	7%	13%	22%	9%	20%	26%
Utilisation index of available time of device work:									
*u*1	0.567	0.499	0.431	0.559	0.492	0.424	0.552	0.484	0.416
*u*2	0.660	0.613	0.565	0.620	0.572	0.525	0.579	0.531	0.484
u3	0.682	0.600	0.519	0.674	0.592	0.511	0.664	0.583	0.501
u4	0.739	0.689	0.639	0.691	0.641	0.590	0.642	0.592	0.542

9. Conclusions

Warehouses influence the quality of services in supply chain. In case of logistics systems quality can be identified with dependability issues, but their dependability cannot be defined and researched like in case technical systems. Dependability of warehouse determines its ability to ensure the continuity of core processes: the production and consumption. This ability can be expressed by a variety of characteristics.

Dependability in the implementation of warehouse processes may be enhanced not only by organisational means, but also technical modifications introducing necessary redundancy to reduce adverse events related to pile-ups in material flows and errors in process realization. Organisational tasks are oriented at increasing the utilisation level of the working time of resources and lowering pile-ups by spreading them over the longer time, while technical modifications are for increasing productivity.

Both actions are intended to increase the probability of correct execution of logistic services by the warehouse, and thus by the entire supply chain. This probability is defined by the *OTIFEF* index. The construction of *OTIFEF* index proposed in this paper is based on the probability of the three basic qualities of a well-executed logistics service. Therefore, it is universal and allows synthetic approach to issues of warehouse dependability referred to the quality of its work. It was pointed out that due to the complexity of operations in supply chains and randomness in the structure and size of material flows, these probabilities can be increased mainly by introducing rational redundancy at the designing stage.

Technical redundancy can be expressed by prosed technical and economic measures, which define frames of dependability structure of warehousing system. This structure can be a base of warehouses and supply chains dependability assessment, but the actual assessment can be made only by analysing historical data of processes realization.

Methods proposed in this paper are applicable. They were developed as elements of SIMMAG 3D project.

Acknowledgement

The research work supported by the National Center for Research and Development, in the frame of PBS 3 project "System for modeling and 3D visualization of storage facilities" (SIMMAG3D).

References

1. Baghalian A, Rezapour S, Farahani R Z. Robust supply chain network design with service level against disruptions and demand uncertainties: A real-life case. European Journal of Operational Research 2013; 227(1): 199–215, https://doi.org/10.1016/j.ejor.2012.12.017.
2. Barnes E, Dai J, Deng S, Down D, Goh M, Lau H C, Sharafali M. On the Strategy of Shupply Hubs for Cost Reduction and Responsiveness. White Paper, The Logistics Institute – Asia Pacific. National University of Singapore, 2003.
3. Bramel J, Simchi-Levi D. The Logic of Logistics: Theory, Algorithms and Applications for Logistics Management. New York: Springer-Verlag, 1997, https://doi.org/10.1007/978-1-4684-9309-2.
4. Bukowski L. System of systems dependability – Theoretical models and applications examples. Reliability Engineering & System Safety 2016; 151: 76–92, https://doi.org/10.1016/j.ress.2015.10.014.

5. Bukowski L A. Zapewnienie ciągłości dostaw w zmiennym i niepewnym otoczeniu. Dąbrowa Górnicza: WSB, 2016.
6. Bukowski L, Feliks J. A unified model of systems dependability and process continuity for complex supply chains. Safety and Reliability: Methodology and Applications / Nowakowski T. [et al.] (ed.). CRC Press Taylor & Francis Group, 2015: 2395-2403.
7. Chung S H, Chan H K, Chan F T S. A modified genetic algorithm for maximizing handling reliability and recyclability of distribution centers. Expert Systems with Applications 2013; 40(18): 7588–7595, https://doi.org/10.1016/j.eswa.2013.07.056.
8. Daganzo C F. Logistics Systems Analysis. New York: Springer Verlag, 1996, https://doi.org/10.1007/978-3-662-03196-4.
9. Grigoroudis E, Siskos Y. A survey of customer satisfaction barometers: Some results from the transportation-communications sector. European Journal of Operational Research 2004; 152: 334–353, https://doi.org/10.1016/S0377-2217(03)00028-6.
10. Haj Shirmohammadi A. Programming maintenance and repair. Technical management in industry, 8th edition. Esfahan: Ghazal Publishers, 2002.
11. Jacyna-Gołda I. Evaluation of operational reliability of the supply chain in terms of the control and management of logistics processes. Safety and Reliability: Methodology and Applications / Nowakowski T. [et al.] (ed.). CRC Press Taylor & Francis Group, 2015: 549-558.
12. Juściński S, Piekarski W. An analysis of a supply process of spare parts for agricultural tractors and machines based on logistic services outsourcing. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2009; 2(42): 63–70.
13. Lewczuk K, Ambroziak T, Warehousing process scheduling in warehouse efficiency and reliability assessment. Proceedings of the 19th International Scientific Conference on Transport Means. Kaunas Univ Technol, Kaunas, 2015: 17–26.
14. Lewczuk K. Dependability issues in designing warehouse facilities and their functional areas. Journal of KONBiN 2016; 2(38): 201–228.
15. Neo H Y, Xie M, Tsui K L. Service quality analysis: case study of a 3PL company. International Journal of Logistics Systems and Management 2004; 1(1): 64–80, https://doi.org/10.1504/IJLSM.2004.005539.
16. Nowakowki T, Werbińska S. Zagadnienie oceny gotowości systemu logistycznego, Logistyka 2007; 5.
17. Nowakowski T. Analysis of possibilities of logistics systems reliability assessment. Safety and Reliability for managing risk 2006; 3. Leiden: Taylor and Francis, 2006.
18. Nowakowski T. Models of uncertainty of operation and maintenance information. Zagadnienia Eksploatacji Maszyn 2000; 35(2): 143–150.
19. Nowakowski T. Niezawodność systemów logistycznych. Wrocław: OWPW, 2011.
20. Nowakowski T. Reliability model of combined transportation system. Probabilistic Safety Assessment and Management. Spitzer C, Schmocker U, Dang V N (ed.). London: Springer, 2004, https://doi.org/10.1007/978-0-85729-410-4_323.
21. Peng P, Snyder L V, Lim A, Liu Z L, Reliable logistics networks design with facility disruptions. Transportation Research Part B-Methodological 2011; 45(8): 1190–1211.
22. Quigley J, Walls L. Trading reliability targets within a supply chain using Shapley's value. Reliability Engineering & System Safety 2007; 92(10): 1448–1457, https://doi.org/10.1016/j.ress.2006.09.019.
23. Rizzi A, Zamboni R. Efficiency improvement in manual warehouses through ERP systems implementation and redesign of the logistics processes. Logistics Information Management 1999; 12(5): 367 – 377, https://doi.org/10.1108/09576059910295805.
24. Rutkowski K. Logistyka dystrybucji. Warszawa: Wydawnictwa SGH, 2009.
25. Santoso T, Ahmed S, Goetschalck M, Shapiro A. A stochastic programming approach for supply chain network design under uncertainty. European Journal Of Operational Research 2005; 167(1): 96-115, https://doi.org/10.1016/j.ejor.2004.01.046.
26. Sawicki P. Wielokryterialna optymalizacja procesów w transporcie. Radom: Wydawnictwo ITE, 2013.
27. Seidler J A. Fundamental concepts of intelligent info system theory, Information Systems Architecture and Technology ISAT '94. Proceedings of 16th Scientific School. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 1994.
28. Sohn S Y, Choi I S, Fuzzy QFD for supply chain management with reliability consideration. Reliability Engineering & System Safety 2001: 72(3): 327–334,https://doi.org/10.1016/S0951-8320(01)00022-9.
29. Starowicz W. Jakość przewozów w miejskim transporcie zbiorowym. Kraków: Wydawnictwo Politechniki Krakowskiej, 2008.
30. Świderski A. Modelowanie oceny jakości usług transportowych. Prace Naukowe Politechniki Warszawskiej. Transport 2011; 81.
31. Wasiak M. Simulation model of logistic system. Archives of Transport 2009; 21(3-4): 189-206.
32. Werbińska-Wojciechowska S. The availability model of logistic support system with time redundancy. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2007; 3(35): 23–29.
33. Werbińska-Wojciechowska S. Time resource problem in logistics systems dependability modelling. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15(4): 427–433.
34. Zamojski W. Teoria i technika niezawodności. Wrocław, 1976.

Ilona JACYNA-GOŁDA
Faculty of Production Engineering
Warsaw University of Technology
Narbutta 85, 02-524 Warsaw, Poland

Konrad LEWCZUK
Faculty of Transport
Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw, Poland

E-mail: jacyna.golda@gmail.com, kle@wt.pw.edu.pl

Article citation info:
GHANOONI-BAGHA M, SHAYANFAR MA, REZA-ZADEH O, ZABIHI-SAMANI M. The effect of materials on the reliability of reinforced concrete beams in normal and intense corrosions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 393–402, http://dx.doi.org/10.17531/ein.2017.3.10.

Mohammad GHANOONI-BAGHA
Mohsen Ali SHAYANFAR
Omid REZA-ZADEH
Masoud ZABIHI-SAMANI

THE EFFECT OF MATERIALS ON THE RELIABILITY OF REINFORCED CONCRETE BEAMS IN NORMAL AND INTENSE CORROSIONS

WPŁYW STOSOWANYCH MATERIAŁÓW NA NIEZAWODNOŚĆ BELEK ŻELBETOWYCH W WARUNKACH NORMALNEJ I SILNEJ KOROZJI

Concrete structures are exposed to a variety of damages during their lifetime each of which could contribute to reducing their service life and load bearing capacity. Since most of parameters have special role in estimating capacity of members which are not certain, probabilistic evaluating the performance of concrete structures could bring more realistic perception about analysis and design of these structures. One of the most frequent probable damages is corrosion. The main focus of this study is placed on reliability assessment of flexural behavior of a reinforced concrete beam experienced pitting corrosion via Monte Carlo simulation. In addition, the effects of time to corrosion initiation, steel rebar diameter, yielding stress of rebars, strength class of cement, aggregate type and compressive strength of concrete, are included both in intense and normal pitting corrosion. The results clearly illustrate that occurrence of intense corrosion in concrete with low compressive strength, which used of higher strength class of cement and crushed stone aggregate, and less initial time for corrosion will lead to considerable reduction in service life even in some cases nearly half.

***Keywords**: failure probability, Monte Carlo simulation, pitting corrosion, flexural capacity, RC beams.*

W trakcie cyklu życia, konstrukcje betonowe są narażone na wiele uszkodzeń, z których każde może przyczyniać się do skrócenia ich żywotności i nośności. Ponieważ większość parametrów odgrywających szczególną rolę w szacowaniu nośności elementów cechuje niepewność, ocena probabilistyczna charakterystyk struktur betonowych może dawać bardziej realistyczny obraz analizy i projektowania tych struktur. Jednym z najczęściej występujących uszkodzeń struktur żelbetowych jest korozja. Głównym celem niniejszego badania była ocena niezawodności zachowania zginanej belki żelbetowej doświadczalnie poddanej korozji wżerowej poprzez symulację Monte Carlo. Ponadto, badano oddziaływanie czasu inkubacji korozji, średnicy stalowych prętów zbrojeniowych, granicy plastyczności tych prętów, klasy wytrzymałości cementu, rodzaju kruszywa i wytrzymałości na ściskanie betonu zarówno w warunkach silnej jak i normalnej korozji wżerowej. Wyniki jasno pokazują, że wystąpienie silnej korozji w betonie o małej wytrzymałości na ściskanie, do produkcji którego wykorzystano cement i kruszywo kamienne o wyższej klasie wytrzymałości, oraz krótszy czas inkubacji korozji prowadzą do znacznego skrócenia żywotności belek, w niektórych przypadkach nawet prawie o połowę.

***Słowa kluczowe**: prawdopodobieństwo awarii, symulacja Monte Carlo, korozja wżerowa, nośność na zginanie, belki żelbetowe.*

1. Introduction

Recent investigations on structures located in coastal areas with hot weather, indicates that sizable number of these structures are at stake of either sulfates or corrosive substances attack [16]. Today, most of impairments during serviceability of concrete structures are concerned to steel corrosion especially in marine structures and concrete bridges [1]. Reinforcing steel is normally passive in concrete due to high alkalinity of the concrete pore solution, however, chlorides or carbonation penetration into the concrete, destroys this inhibitive property of the concrete and leads to corrosion [38]. Two types of corrosion –general and pitting –are possible [14]. General corrosion affects on cross section of reinforcement with more or less uniform metal loss over the perimeter of reinforcing bars [28]. It also causes cracking and eventually spalling of the concrete cover and produces rust staining on the concrete surface so it can be detected quite easily during inspection of a structure [39]. Pitting or localized corrosion, in contrast to general corrosion, concentrates over small areas of reinforcement. Pitting corrosion often does not cause disruption of the concrete cover and produces little rust staining on the concrete surface and so it may be difficult to be discovered during inspection [18-19]. Despite that a numerous researches have been done about time dependent reliability of these deteriorating structures, most of them are based on uniform corrosion assumption [12, 15-16].

Stewart [34], Stewart & Al Harthy [32] and Stewart & Suo [35] studied the spatial effects of local corrosion in reliability of bending behavior of concrete beams. These researchers developed a Stochastic model for pitting corrosion in a simply supported beam in which Eigen value theory is applied with the intention of predicting the maximum value of corrosion depth as function of both diameter and length of steel rebar. Note that, they presumed following assumptions: i) statistical data about corrosion depth were gathered based on limited number of previous studies, ii) ductile rapture mode of steel rebar is presumed, iii) the effect of corrosion on steel material is neglected.

Subsequently Stewart & Al Harthy [32] elevated previous study [34 , 33] by employing a progressive corrosion data on two rebar diameter namely 16 and 27 in order to achieve modified Eigen value for pitting corrosion. Unlike the former study, they also considered the corrosion impact on steel material. Interestingly in the present article, the same approach is recruited, will be depicted in Sec.4.Barone and Frangopol investigated the probabilistic approaches and obtained that Maintenance schedules are for life-cycle maintenance of structural systems [3]. Zhang et al. presented the time-dependent probability of chloride-induced corrosion of reinforced concrete (RC) structures in marine environments, by developing the third-moment (TM) method [41].According to Simioni [31] and Pedeferri [27], often the amount of cross-sectional reduction is intense and neglecting it in the design of RC structures can lead to serious damage.

Based on the studies conducted above, the outcome of section area reduction due to corrosion intense is not discussed; so, in the present article, a probabilistic model for flexural behavior of a beam is suggested in pursuit of reliability assessment of concrete beams with rectangular section which are subjected to pitting corrosion with two kinds of pitting scenarios (intense and normal pitting) based on the amounts of pit depths. More precisely, all the possible detrimental impacts of corrosion comprised of cross section reduction, decrease in yield stress and so on, are considered in the process. Based on suggested probabilistic model, reliability assessment of a reinforced concrete beam will be evaluated via employing Monte Carlo simulation. Finally, according to the results, the role of different parameters in serviceability of beam will be reported.

2. Deterioration process

After initial corrosion and during of corrosion propagation, load bearing capacity of RC members decreased [28].

2.1. Corrosion initiation

The time to corrosion initiation will be calculated regard to second diffusion law in 1 dimensional in semi-finite solid body (Fick's law according to Eq. (1)), shown in Eq.(2). It should be noted that diffusion in the practical process are different from that assumed with Fick's law, yet it is used normally for simulating chloride ingress in concrete, because of its accuracy [5, 2]:

$$(\frac{\partial f}{\partial t} = D_c \frac{\partial^2 f}{\partial x^2}) \tag{1}$$

$$T_0 = \frac{C^2}{4D_c}[erf^{-1}(1-\frac{C_{th}}{C_0})]^{-2} \tag{2}$$

where "C" is the cover (cm), "erf" indicate the error function, C_{th} and C_0 is the chloride threshold and constant amount of chloride on surface (kg/m^3) respectively and D_c is indication of chloride diffusion coefficient (cm^2/s), which directly correlated with Water-cement ratio, environmental condition like temperature and humidity level. It will be calculated as follow [4, 6]:

$$D_c = [1.249 - 5.051wc + 8.941wc^2] \times 10^{-8} \tag{3}$$

According to the building research establishment of Iran [7], the relationship between the water to cement ratio (W/C), the type of aggregate in the concrete and strength class of cement with compressive strength of concrete are obtained by the following relationships.

For crushed aggregate and strength class of cement(C 525):

$$F_c(MPa) = 573.91(w/c)^3 - 984.34(w/c)^2 + 435.71(w/c) + 4.4767 \tag{4-a}$$

For round aggregate and strength class of cement (R 525):

$$F_c(MPa) = 545.45(w/c)^3 - 894.81(w/c)^2 + 369.48(w/c) + 13.818 \tag{4-b}$$

For crushed aggregate and strength class of cement (C425):

$$F_c(MPa) = 525.25(w/c)^3 - 817.75(w/c)^2 + 313.6(w/c) + 22.121 \tag{4-c}$$

For round aggregate and strength class of cement (R 425):

$$F_c(MPa) = 707.07(w/c)^3 - 1087.4(w/c)^2 + 450.57(w/c) + 5.6061 \tag{4-d}$$

2.2. Corrosion propagation

Quantitative definition of corrosion propagation is provided primarily in the form of corrosion rate which is offered as amount of steel loss per unit time and unit area. Most of current non-destructive techniques detects the corrosion whereby electromechanical measurement of current rate (i_{corr}).Stewart et al, developed an analytical model for estimating the corrosion which is dependent upon parameters like quality of concrete and rebar cover [17, 39]:

$$(i_{corr}(T) = 0.85 i_{corr}(1).T^{0.29}) \tag{5}$$

In which "T" indicates how long corrosion have initiated ($T = t - T_0$), T0 shows the initiation time of corrosion and "t" is the structure life (year). This equation is viable to be used for cities located in Asia, Europe, America and Australia [39, 40]. One major drawback of this equation is affiliated with corrosion with short periods of structure life. In other words, if "T" is small $(t \approx T_0 \rightarrow T \approx 0)$, corrosion rate tends to reach infinity. In this respect, modified mentioned equation as below [40]:

$$i_{corr}(T) = \begin{cases} i_{corr}(1) & 0 \le T \le 1\ year \\ 0.85 i_{corr}(1).T^{-0.29} & T > 1\ year \end{cases} \tag{6}$$

In which $i_{corr}(1)$ is the one year corrosion rate at the start of corrosion initiation $(\mu A / cm^2)$, which is suggested as following equation [10]:

$$i_{corr}(1) = \frac{37.8(1-wc)^{-1.64}}{C} \tag{7}$$

In general corrosion in order to calculate corrosion rate, according to Faraday's law of electrochemical equivalence will be recruited into steel mass loss, which i_{corr} the corrosion rate. ΔD indicates the diameter reduction (mm) and "T" illustrate the number of years [8]:

$$\Delta D(T) = 0.0232 i_{corr} * T \tag{8}$$

$$A_s(T) = n\frac{\pi[D_0 - \Delta D(T)]^2}{4} \geq 0 \tag{9}$$

If corrosion current density is assumed to be identical for "n" steel rebars with diameter D_0, the cross section area of rebars could be resulted according to Eq.(9), but if corrosion is local, the cross section of the rebars should be calculated as:

$$A_s(T) = n\left(\frac{\pi D_0^2}{4} - A_{Pit}\right) \geq 0 \tag{10}$$

where $A_{Pit}(t)$ and A_s is the cross section area of corroded and non-corroded rebar, respectively. "n" is cited as the number of rebar, and $A_{Pit}(t)$ will be seen in Fig. 1 and calculated as following equation [38, 35, 21]:

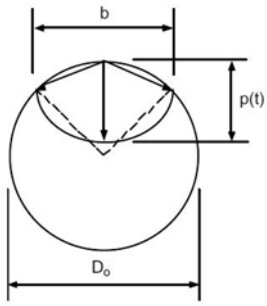


Fig. 1. Configuration of pitting corrosion in rebar [38, 41]

$$A_{Pit}(T) = \begin{cases} A_1 + A_2 & P(T) \leq \frac{D_0}{\sqrt{2}} \\ A_s - A_1 + A_2 & \frac{D_0}{\sqrt{2}} < P(T) \leq D_0 \\ A_s & P(T) > D_0 \end{cases} \tag{11}$$

where:

$$P(T) = 0.0116.i_{corr}(T).R.T \tag{12}$$

$$A_s = \frac{\pi.D_0^2}{4} \tag{13}$$

$$A_1 = 0.5[\theta_1(\frac{D_0}{2})^2 - b\left|\frac{D_0}{2} - \frac{P(T)^2}{D_0}\right|] \tag{14}$$

$$A_2 = 0.5[\theta_2 P(T)^2 - b\frac{P(T)^2}{D_0}] \tag{15}$$

$$b = 2P(T)\sqrt{1-(\frac{P(T)}{D_0})^2}\ ,\theta_1 = 2\arcsin(\frac{b}{D_0})\ ,\theta_2 = 2\arcsin(\frac{b}{2P(T)}) \tag{16}$$

In above equation, $P(T)$ is the maximum pit depth (mm), R is pitting factor will calculated as the ratio of $P(T)$ to average corrosion in a period of years $(R = P(T)/P_{av})$. Additionally, $i_{corr}(T)$ is corrosion rate, which is a function of time (in term of $(\mu A/cm^2)$), and T were defined in the previous section. In fact, in case of normal corrosion with usual penetration rate of flexural capacity reduction of beam as a result of the loss of transverse cross-sectional area of rebars (according to Eqs. 12-16), Even after 100 years it is negligible and with fact of many observations of marine structures in extreme areas, is inconsistent. For this reason, in this study, to achieve a more realistic behavior of structures under severe corrosion, pitting depth in Eq. (12), was considered ten times the usual amount.

3. Reliability assessment

Reliability of many practical problems can be investigated through a mathematical model composed of "R" related to strength of material and "S" associated with imposed load. These mathematical models are known as limit state functions indicative of an intended behavior of the structure. It is worth to mention that a damage will be induced to the structure when the imposed load are larger than structure capacity; that is, when $G = R - S$ is minus [21, 23]:

$$P_f = \iint_{R-S<0} f_{S,R}(s, r)dsdr = P[G(U) < 0] = \int_{G(U)<0} f_X(X)dX \tag{17}$$

in which P_f, $G(U)$ and $f_X(X)$, are failure probability of structure, limit state function and joint probability density function of variables, respectively.

The major problem in reliability assessment of structures is solution of abovementioned integration. One of the suggested approximate solutions for mentioned integration is Mont Carlo simulation [30, 23] .According to the definition counter function, failure probability can be written as [23]:

$$\begin{aligned} P_f &= \int_{G(x)\leq 0} f_{x_1\cdots x_n}(x_1, \cdots, x_n)\mathrm{d}x_1\cdots dx_n \\ &= \int_{\infty}^{\infty}\cdots\int_{-\infty}^{\infty} I[X]f_{x_1\cdots x_n}(x_1, \cdots, x_n)dx_1\cdots dx_n \end{aligned} \tag{18}$$

And finally P_f can be written as follow:

$$P_f = \int_{G(x)\leq 0}\cdots\int f_{x_1\cdots x_n}(x_1, \cdots, x_n)dx_1\cdots dx_n \approx \frac{1}{N}\sum_{i=1}^{N} I[X_i] \tag{19}$$

where:

$$I[X] = \begin{cases} 0 & G(x) > 0 \\ 1 & G(x) \leq 0 \end{cases} \tag{20}$$

4. Probabilistic beam model for corrosion

A probabilistic model is comprised of a variety of random variables such as load and material strength which will be analyzed via a limit state function. This model should be inclusive of both all possible uncertainties in reliability assessment and statistical parameters of the damage. Consequently, at first in this section devoted to propose the desirable limit state function suited for flexural behavior of reinforced concrete beams, at last aimed at offering felicitous statistical specifications of time to corrosion initiation formula.

4.1. Limit state function

Nominal ultimate flexural strength of a concrete beam with rectangular section could be readily resulted from bending theory. By considering characteristic and equilibrium equations, nominal flexural capacity will yielded (M_n) as it is explicit observable from the Eq.(21), there is positive correlation between area of steel rebars and flexural strength [35, 4]:

$$M_n = A_s(t) f_y \left[d - \frac{A_s(t) f_y}{1.7 b f_c'} \right] \tag{21}$$

According to mention in sec 3, limit state function for flexural capacity written as follow;

$$G(M) = M_n - M_u \tag{22}$$

Where M_u is imposed moment due to external loads and (M_n) is nominal flexural strength of the beam. According to above equation, the probabilistic model could be developed as following equation involving both uncertainties of strength and steel deterioration:

$$G(M) = ME_m A(T) f_y(T)(d - K\frac{A_s . f_y(T)}{b.f_c}) - M_u \tag{23}$$

$$fy(T) = (1 - \alpha \frac{A_s - A_s(T)}{A_s}) \mathrm{f}_y \tag{24}$$

In the above limit state function, $f_y(T)$ is yield stress of corroded steel which is linearly correlated with un-corroded one, $(A_s - A_s(T))$ and α coefficient is an experimental parameters which is normally recommended to be 0.5 [35, 21]. A numerical study with full detail is available in ref. [21] which properly predicts the capacity of the structures subjected to corrosion, yet even these analysis do not consider the spatial effects of corrosion and redistribution of load among residual rebars after at least a failure. These fields of studies requires more research in future. It is instructive to note that in the limit state function mentioned in Eq.(23),uncertainties were considered in forms of ME_m [35, 26, 23].

5. Reliability assessment of corroded RC beam

Limit state function mentioned in Eq.(23), is a probabilistic formulation for prediction of occurring flexural failure of a concrete beam, which also embraces independent random variables. These variables are simulated by a group of probability functions such as Normal, Log-normal and Gumbel distribution function. Note that varying properties of corrosion initiation is resulted from statistical analysis and the rest of properties are assumed based on previous studies in literature. Accordingly, the analytical model is composed of a one span concrete beam with simply supports at the ends subjected to a uniform distributed load. It has 10 m length, and rectangular cross-section with 0.35 width and 0.8 height. Two groups of rebar layout are considered for longitudinal reinforcing steel one of which has 9 rebars with 25.4 mm diameter (A_s =4560 mm^2) and the other one possesses 3 rebars with 43 mm diameter(A_s =4356 mm^2). Finally, as mentioned before, independent variables in limit state function will be analyzed via Monte Carlo simulation.

5.1. Statistical properties of corrosion variables

As it shown in Eq.(6), corrosion rate is relevant upon concrete characteristics and initiation of corrosion. Besides, according to Eq. (2), initiation time, itself, is dependent on concrete cover, chloride threshold and constant amount of surface chloride and diffusion coefficient. Similarly, diffusion properties is also at the mercy of water to cement ratio; hence, it can be concluded that initiation time of corrosion is a function of water to cement ratio, that is related to compressive strength of concrete.

In Tables 1-3 each random variables with mean and coefficient of variation and type of distribution are defined.

5.2. Pitting factor (R)

As it was mentioned earlier, unlike general corrosion, pitting corrosion aims at limited area of rebars; thus, the corroded area in pitting corrosion might be much less than corrosion resulted from i_{corr} . As a result, i_{corr} should not be employed directly to evaluate the cross-sectional corroded area. According to study conducted by Gonzaleset al. [16], the maximum pit depth is roughly 4-8 times bigger than correspondent value for general one. For example, results for rebars with length of 125 mm and diameter of 8 mm are in accordance with results in Tuutti1982 [37] in which "R" for rebars with length of 150 – 300 mm and diameter bounded below by 5 mm and above by 10 mm is reported between 4-10.

Based on Tuutti 1982 [37], there is a considerable uncertainty for parameter "R". In order to address this issue, one of the most common approaches is probabilistic modeling via eigen value theory. For this purpose, this parameter is modeled as random variable Gumbel distribution, resulted form Turnbull 1993 [36]. Note that since this approach is pervasively recruited for steel plates, pipelines and prestressed cables, which leads to acceptable results, it is reasonable to employ that here to involve uncertainty for "R".

Table 1. Statistical characteristics of effective variables in estimating the corrosion initiation

Parameter	Mean	COV	Distribution	Reference
Concrete cover (mm)	50	0.12	Normal	Darmawan et al.(2010) [11]
c_0 ($\frac{kg}{m^3}$)	3.05	0.74	Normal	Stewart (2004) [34]
c_{th} ($\frac{kg}{m^3}$)	0.9	0.2	Log-normal	Stewart (2004) [34]
D_c ($\frac{mm^2}{year}$) (High quality)	14	0.75	Log-normal	Nogueira &, Leonel (2013) [22, 29]
D_c ($\frac{mm^2}{year}$) (Low quality)	163	0.75	Log-normal	Nogueira &, Leonel (2013) [22, 29]

Table 2. Statistical characteristics of uncertainty variables in model

Parameter	Distribution	COV	Mean	Reference
ME_m (flexural model uncertainty)	Normal	0.120	1.10	Nowak et al(2005) [24]
K	Normal	0.05	0.59	Stewart (2009) [35]
α	Log-normal	0.12	0.5	Stewart (2004) [34]
Parameter of pitting corrosion (R)	Gumbel	0.22	5.65	Stewart (2004) [34]

Table 3. Statistical characteristics of mechanical features of materials, load and resistance

Parameter	Mean	COV	Distribution	Reference
f_c concrete compressive strength (MPa)	27.6	0.18	Log-normal	Mirza et al.(1979) [20]
f_y steel strength (MPa)	414	0.10	Log-normal	Mirza et al (1979)[20]
Dead load($\frac{KN}{m}$)	$1.05G_n$	0.10	normal	Ellingwood et al.(1980) [13]
Live load($\frac{KN}{m}$)	$0.3Q_n$	0.6	gamma	Chalk and Corotis (1980) [9]
d(effective depth of cross section) (cm)	nominal	0.02	Normal	Ostlund(1991),Lu et al. (1994)[25]
Concrete cover (cm)	9	0.12	Normal	Darmawanet al.(2010)[11]

5.3. Pitting factor via Gumbel distribution

The following equation is used to compute R as a random variable modeled by the Gumbel distribution:

$$F(R) = \exp\left\{-\exp\left[-\frac{(R-\mu)}{\alpha}\right]\right\} \tag{25}$$

Where μ and α are Gumbel distribution parameters. According to Stewart 2004 [34], R parameter varies from 4 to 8with assumption for rebar with 8 mm diameter and 125 mm length (between of 5% to 95% in Gumbel distribution).Then, mean value and C.O.V are 5.65 and 0.22, respectively, which is related to μ_0=5.08 and α_0=1.02. In order to compute Gumbel distribution parameters for other rebars with different diameter, ensuing equation is suggested by Gonzales et al. [16].

$$\alpha = \alpha_0 \quad \mu = \mu_0 + \frac{1}{\alpha_0}\ln\left(\frac{A}{A_0}\right) \tag{26}$$

Where "A" is lateral area of the rebar and A_0 is lateral area of rebar with 8 mm diameter and 125 mm length. There are also evidences indicating that pitting factor might decrease over the time [6].Thus, statistics reported for pit depth ought to be used conservatively.

6. Analysis Results

In general and base on mentioned equations, mostly corrosion occurrence in normal pitting corrosion, but sometimes pitting depth is very higher than equation. As it can be observed in Fig.2, pitting corrosion due to intensity of chloride concentration results to a high amount of pit depth. In practice, however, this depends on such parameters like how long and to what extent rebars have been subjected to corrosion. So as to consider this uncertainty in the simulation, corrosion parameters except pitting depth will be defined same for both intense and normal corrosion. For the first case, random variables of P(T) for calculating A_{pit} in Eq.(12), considered about 10 times of normal corrosion that corrosion could progress as much as possible. Accordingly intense corrosion will be defined in Matlab Software leading to the fact that corrosion depth could be extended nearly up to rebar diameter. But for the latter case, normal pitting, Eqs.(11) to (16) will be considered for corrosion parameters, that corrosion would be restricted by less than 0.2 rebar diameter. In this regard, two modeling approach will be recruited.

The concrete beam which mentioned in previous section, is reliability assessed with 3e5 and 3e7 simulations of Monte Carlo for intense and normal pitting corrosion with two layout of reinforcing steels that defined in sec 5. In each analysis, one of the random variables playing role in flexural capacity of beam is considered with the intention of scrutinizing the correspondent changes.

6.1. Effect of aggregate type and strength class of cement

In this section, the correlation between strength class of cement and failure probability will be discussed. Figs. 3 and 4 Portrays the effect aggregate type(R=Round aggregate and C=Crushed aggregate) and strength class of cement(425 kg/cm^2 and 525 kg/cm^2)for two layouts of longitudinal rebars.

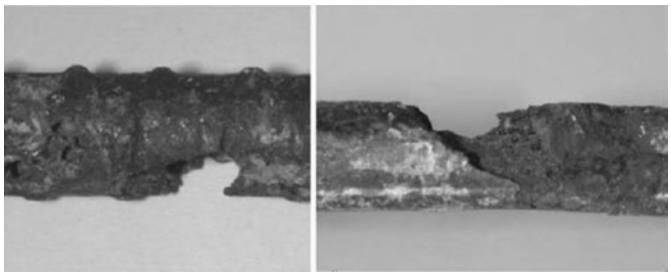


Fig. 2. The effect of intense pitting corrosion on rebar [27]

According to the presented formula in relationship 4, for fixed compressive strength $f_c = 27.6\ MPa$, for each strength class of cement and the applied aggregate, it is possible to determine the ratio of water to cement to achieve the resistance of interest. Also, according to relationships 6 and 7, after determining the corrosion rate as decreased armature area based on 11-16, it will be possible to obtain decreased flexural capacity of RC beams. By comparing the diagrams of the following charts, it will be possible to observe the effects of various strength class of cement as well as the effect of aggregate shape on the failure caused by corrosion.

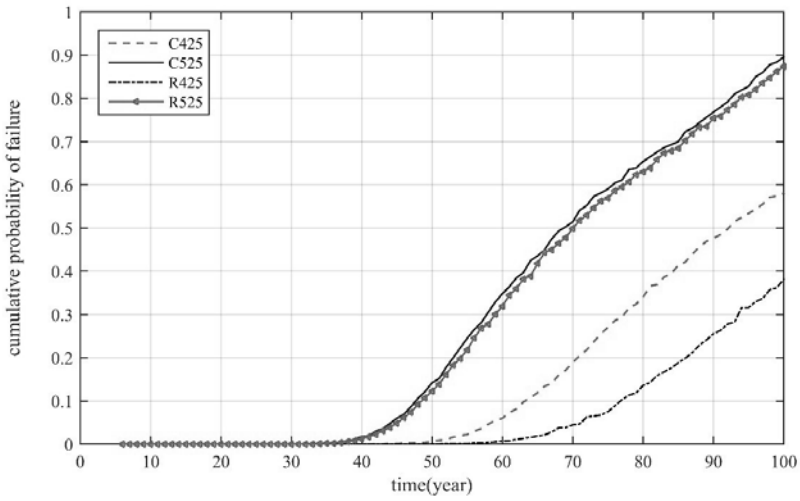


Fig. 3. Beam with rebars diameter of 25.4 mm with different cement strength class and aggregate type

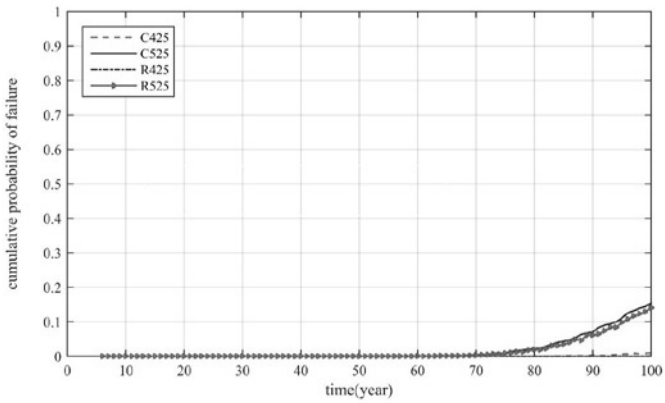


Fig. 4. Beam with rebars diameter of 43 mm with cement strength class and aggregate type

Despite of expectations in mix design, in case that round shaped aggregate isused for concrete mixture, the concrete beam is less probable to fail during its lifetime. In other words, usage of round shaped gravel will postpone the failure time. It is worth to mention that this delay is more considerable when it comes to concrete with smaller strength class of cement. For example, for C425, postponement is around 10 years whereas for C525 is nearly one year.

It is worth mentioning that regarding relationship (4) (f'c-wc), for a constant compressive strength of concrete as the strength class of cement is higher, higher water to cement ratio will be obtained. As a result, as expected, with increased ratio of water to cement (in fixed resistance), the conditions for the formation of corrosion will be prepared and the failure probability increases. That is why, according to Figs.3 and 4, when we use cement with higher strength class, the failure probability increases.

Of course, the use of various cement strength class can implement two reverse effects on the failure probability of concrete beam; in fixed water to cement ratio, if cement with higher strength class is used, higher compressive strength of concrete will be obtained and the failure probability decreases. But in fixed compressive strength of concrete, in the case of using higher resistance, the water to cement ratio increases and the failure probability of concrete beam will be more likely.

6.2. Effect of compressive strength of concrete

The focus of this section is placed on evaluating the role of compressive strength of concrete in reliability assessment of the beams. Figs. 5 and 6 outlines the failure probability of the two mentioned beams with two rebars layout and different compressive strengths.

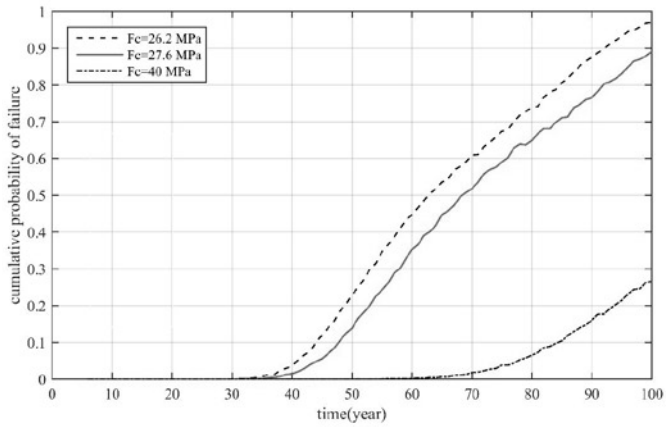


Fig. 5. Beam with rebars diameter of 25.4 mm with different compressive strength of concrete

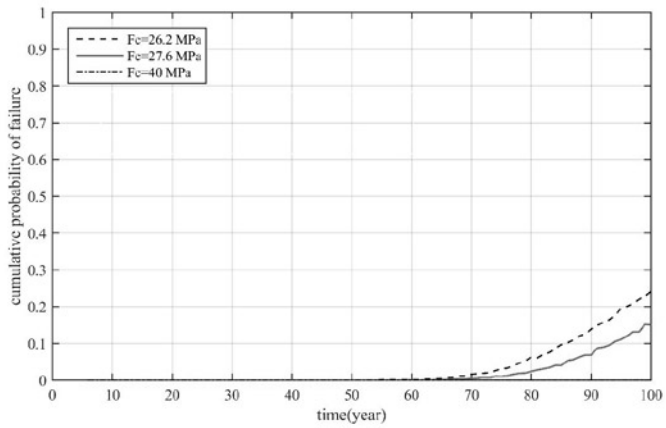


Fig. 6. Beam with rebars diameter of 43 mm anddifferent compressive strength of concrete

As it is observable in both figures, as the compressive strength grow, concrete beams are more resistant to corrosion which itself will lead to declining the failure probability.

6.3. Effect of yield stress of steel rebars

As the fourth factor that evaluated, the yield stress of reinforcing steel performing a role in reliability assessment of the concrete beams and is investigated here in this section. The effect of yield stress of rebars are presented in Figs. 7 and 8 in which are for beams reinforced with 25.4 mm and 43 mm diameter, respectively. As it is obvious in both figures, augment in yield stress will make structure

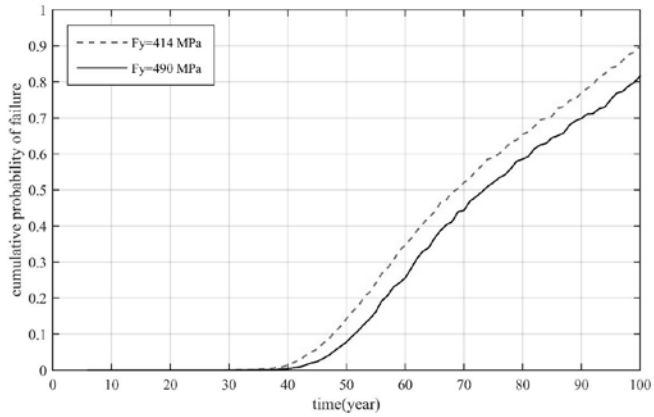


Fig. 7. Beam with rebars diameter of 25.4 mm and different yield stress

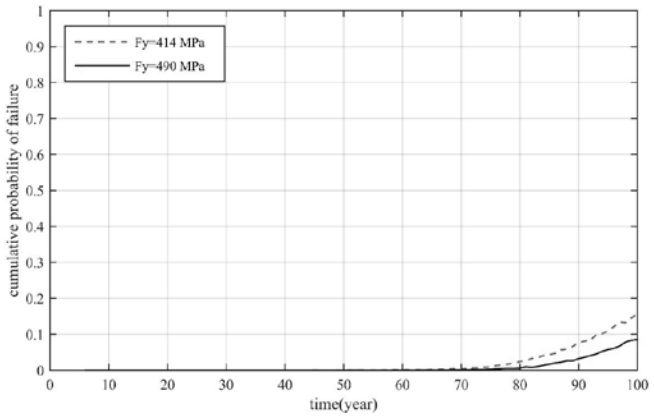


Fig. 8. Beam with rebars diameter of 43 mm and different yield stress

more immune to corrosion, but it should be noted that this change is relatively low.

6.4. Effect of diameter of steel rebars

The last but not the least parameter played undeniable role in reliability assessment of the concrete beams is pertinent to arranging steel through out the cross section. More particularly, if "A_s" cm^2 is required to reinforce the section, how many and what diameters are needed. The two following figure will enlighten the role of rebar diameter for identical area of steel. In two beams with the identical physical and mechanical properties, if the diameter of longitudinal rebar is 43 mm, probability of failure is zero for the first 40 years henceforth it soars up till the 100th year when the probability will see 70 %. However, if longitudinal rebar has the diameter of 25.4 mm, the same pattern happen for the first twenty years, the failure probability is zero, but it will experience an ascending trend after 20th years till the 66th year when the probability is 100 %.

The main reason that rebar with lower diameter after corrosion (assuming similar corrosion initiation time and ignoring cover changes in this parameter) has higher failure probability is that the corrosion penetration depth is similar for both rebar diameters. Therefore, the ratio of decreased area to the initial area and as a result, decreased flexural resistance percentage will increase in rebars with lower diameter. This issue, along with fixed load, leads to higher failure probability.

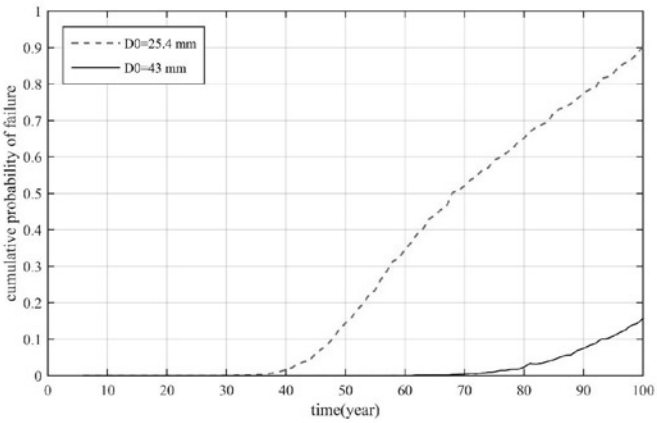


Fig. 9. Beam with different diameter at the same cross-sectional area

It should be note that in the presented diagram in Fig.9, the effects of increased diameter in decreasing cover on rebar have been ignored. Furthermore, in these diagrams, the effects of R variation in rebars with more diameter have been ignored. It should be noted that according to concrete quality and cover, the time difference for the corrosion initiation for rebars with different diameters changes. It is obvious that when the corrosion initiation for both rebars is almost similar, the results of Fig.9 are valid. But increasing the time for the corrosion initiation based on Figs. 5 and 6 changes the results in Fig.9. Also, for higher time difference, the corrosion initiation for different rebar diameters can be reversed.

If the parameters presented in Tables 1& 3 are used for concrete beam that studied in this paper, and for low and high quality concrete we used diffusion coefficients of D_c=163 mm^2/year and D_c=14 mm^2/year, the probability for corrosion for both rebar and concrete cover can be observed in the following chart. It is observed that for high quality concretes (D_c=14), corrosion will not occur for 50 years. For low quality concretes (D_c=163), the time difference for the corrosion initiation time for both rebars with diameters of 25.4 mm and 43 mm will be about 1 year as 7 and 6 years, respectively.

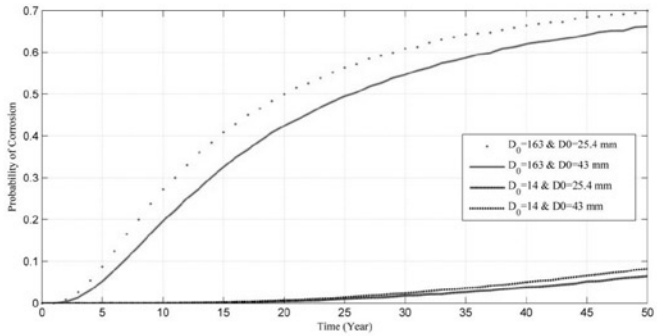


Fig. 10. The corrosion probability for rebars in both type concretes (net cover=90-(D0/2)

In the case of using concrete with the 50 mm cover, as can be seen from Fig. 11, if we consider the corrosion criteria as 10%, regarding concretes with high quality with 50 years lifetime, if rebars with diameters of 25.4 mm and 43 mm are used, the corrosion initiation will be 2 and 1 years, respectively.

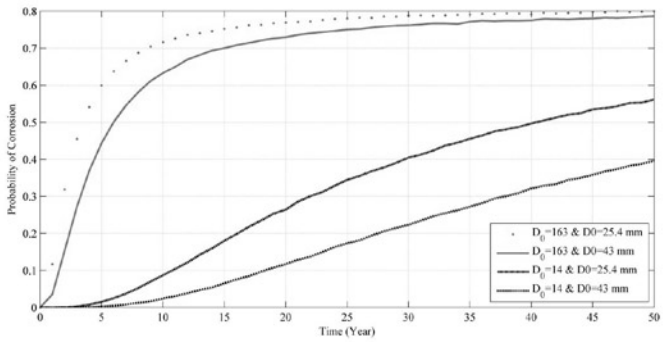


Fig. 11. The corrosion probability for rebars in both type concretes (net cover=50-(D0/2)

According to Figs. 10 and 11, the highest time difference for the corrosion initiation for two different rebars while using high quality concrete (lower water to cement ratio and lower diffusion coefficient), with the concrete cover is lower. Therefore, in this condition, due to the occurrence of rapid corrosion in larger rebars, the damage to the structures with larger rebars will be larger.

6.5. Effects of corrosion initiation time

Figs. 12 and 13 are diagrams depict failure probability of the beam with different time of corrosion initiation and different rebar layout. By comparing figures, the impact of initiation time of corrosion on flexural failure probability will be illuminated. It is observed that as

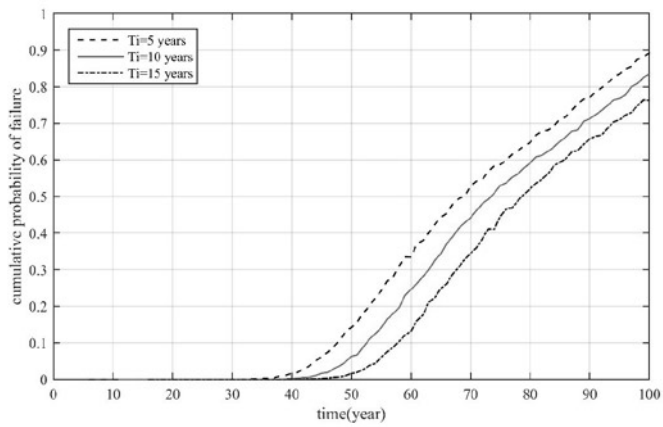


Fig. 12. Beam with rebars diameter of 25.4 mm at different time of corrosion initiation

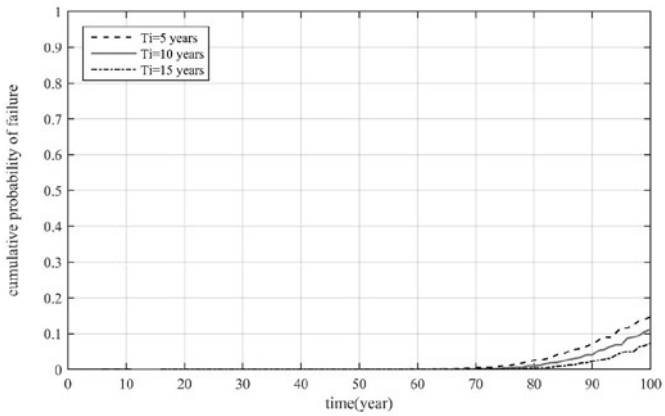


Fig. 13. Beam with rebars diameter of 43 mm at different time of corrosion initiation

the time of corrosion initiation is increased (lower diffusion coefficient and larger cover), the failure probability of structure is lower.

As it can be perceived in Figs. 12 and 13, as the initiation time of corrosion rises, the probability of failure decreases as expected. However, for a given service life of 100 years, it can be observed that the bigger diameter for rebars leads to the less failure probability of the beam. An illustration of this could be seen for the two beams. The samples reinforced with 25.4 mm rebar diameter, all fails are about 75 to 90 % after 100 years while the other group of beam (with 43 mm rebar diameter) are remain less than 20 percent of failure probability after 100 years.

6.6. Corrosion with normal pitting

In this type of pitting the concrete beam mentioned in previous sections, is reliability assessed with 3e7 simulations of Monte Carlo. All of discussed in previous section could be generalized for the case of corrosion with normal pit with exception that in this case all the failure probability will significantly decline in comparison with intense pitting. Besides, according to Figs.14-16, the range of variation in results is inconsiderable for a period of 100 years. So it can be concluded that the whole process of corrosion with normal pit can be neglected. Main reason of in subject, is slight reduce in cross-section according to Eqs.(12) to (16) even after 100 years. That lead to almost constant flexural capacity without considerable reduction and subsequently limit state function remain positive in flexural capacity.

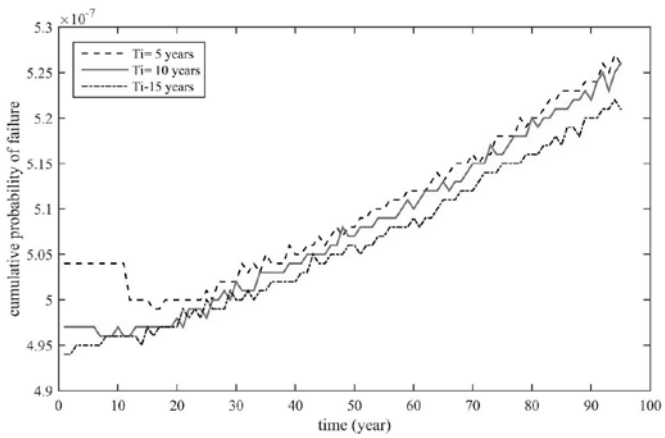


Fig. 14. Beam with rebars with a diameter of 25.4 mm at different corrosion beginning times

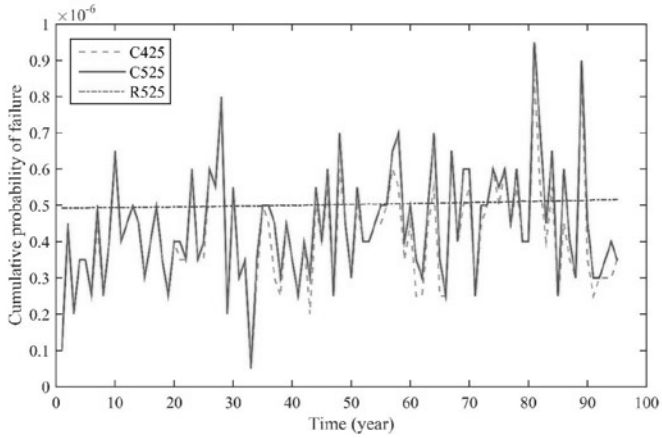


Fig. 15. Beam with rebars diameter of 43 mm with different aggregate and cement strength class

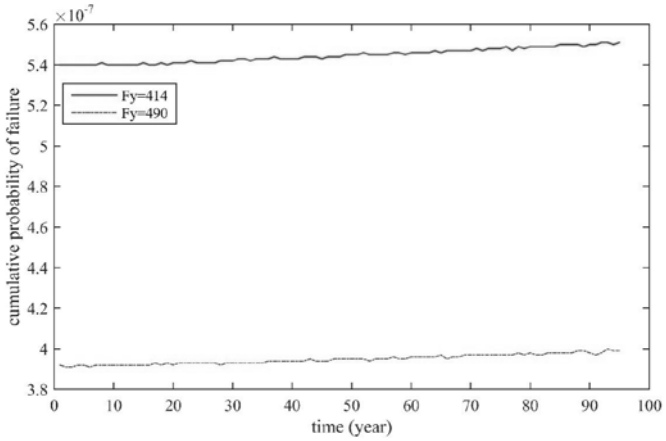


Fig. 16. Beam with rebars with a diameter of 25.4 mm with different beam yield stress

7. Conclusion

In the present study, a probabilistic model for reliability assessment of concrete beams involving various kinds of uncertainties, compressive concrete strength and steel layout, is proposed. Two structures are reliably assessed for residual life-time. Besides, the effect of aggregate type, strength class of cement, compressive strength of concrete, yield stress of rebar and bar diameter are also evaluated. According to performed reliability assessment, amount of water to cement ratio of mixture is the most important parameter for reliability of corroded RC beams. Following points could be summarized for the analysis results:

- Enhancing compressive strength of concrete as well as concrete cover will lead to making the structures more resistant to corrosion. It should be noted that increasing concrete cover is the most inexpensive and practically feasible way to protect the reinforcing steel from corrosion. For example, for the case of diffusion coefficient and D_c=14 mm^2/year, if the cover increase 40 mm, the failure probability might decrease between 30% - 45%.
- Increase of time to corrosion initiation (with less diffusion coefficient of concrete or high quality of concrete and higher concrete cover) contributes to decrease the failure probability of beams.
- According to the results, utilization of aggregate with round shaped corner will adjourn the failure of the structures. This delay in failure of structures is more considerable for concrete with less strength class of cement.

- According to the results, failure probability of corroded beam, in case of beam with higher compressive strength, is less.
- Augmenting the yield stress of longitudinal reinforcement will slightly reduce the failure probability of concrete corroded beams.
- With assumption of equal initiation time of corrosion for rebar with different diameter (in term of effecting parameters for corrosion initiation such concrete cover and diffusion coefficient), usage of thicker reinforcement when the same area of steel, contribute to decline the failure probability of beams. But in case of high quality of concrete mix, with lower concrete cover, usage of thicker reinforcement with the same area of steel, contribute to increase the failure probability of beams.
- - It is observed that in case of local corrosion with normal pitting, the same results will be reported, but the difference is that the failure probabilities are far less than intense corrosion. In this case, since most of the time, safety factor is large, corrosion could be totally neglected as a detrimental factor.

Acknowledgement

The present study was financially supported by Committee for Scientific Research of Islamic Azad University, East Tehran Branch, Tehran, Iran (IAUET).

References

1. Ahmad S. Reinforcement corrosion in concrete structures, its monitoring and service life prediction–A review. Cement and Concrete Composite 2003; 25: 459-471, https://doi.org/10.1016/S0958-9465(02)00086-0.
2. Almusallam A A. Effect of degree of corrosion on the properties of reinforcing steel bars. Construction and Building Materials 2001; 15(8): 361-368, https://doi.org/10.1016/S0950-0618(01)00009-5.
3. Barone G, Frangopol D M. Reliability, risk and lifetime distributions as performance indicators for life-cycle maintenance of deteriorating structures. Reliability Engineering & System Safety 2014; 123: 21-37, https://doi.org/10.1016/j.ress.2013.09.013.
4. Bastidas-Arteaga E, Bressolette P, Chateauneuf A, Sánchez-Silva M. Probabilistic lifetime assessment of RC structures under coupled corrosion–fatigue deterioration processes. Structural Safety 2009; 31(1): 84-96, https://doi.org/10.1016/j.strusafe.2008.04.001.
5. Bastidas-Arteaga E, Sánchez-Silva M, Chateauneuf A, Silva M R. Coupled reliability model of biodeterioration, chloride ingress and cracking for reinforced concrete structures. Structural Safety 2008; 30(2): 110-129, https://doi.org/10.1016/j.strusafe.2006.09.001.
6. Bhargava K, Mori Y, Ghosh A K. Time-dependent reliability of corrosion-affected RC beams—Part 1: Estimation of time-dependent strengths and associated variability. Nuclear Engineering and Design 2011; 241(5): 1371-1384, https://doi.org/10.1016/j.nucengdes.2011.01.005.
7. Building Research Establishment. Iranian Design Code for Normal Concrete Mixes. second edition 2005.
8. Bushman J B, Engineer P P. Calculation of Corrosion Rate from Corrosion Current (Faraday's Law). Bushman & Associates Inc. 2000.
9. Chalk P L, Corotis R B. Probability model for design live loads. Journal of the Structural Division. 1980 Oct; 106(10): 2017-2033.
10. Dai H, Wang W. Application of low-discrepancy sampling method in structural reliability analysis. Structural Safety 2009; 31(1): 55-64, https://doi.org/10.1016/j.strusafe.2008.03.001.
11. Darmawan M S. Pitting corrosion model for reinforced concrete structures in a chloride environment. Magazine of Concrete Research 2010; 62(2): 91-101, https://doi.org/10.1680/macr.2008.62.2.91.
12. Dimitri V V. Deterioration of strength of RC beams due to corrosion and it's influence on beam reliability, Journal of Structural Engineering 2007; 133: 15-42.
13. Ellingwood B. Development of a probability based load criterion for American National Standard A58: Building code requirements for minimum design loads in buildings and other structures. US Department of Commerce, National Bureau of Standards; 1980, https://doi.org/10.6028/nbs.sp.577.
14. Enright M E, Frangopol, D M. Probabilistic analysis of resistance degradation of reinforced concrete bridge beams under corrosion. Engineering Structures Journal 1998; 20(11): 960-971, https://doi.org/10.1016/S0141-0296(97)00190-9.
15. Frangopol D M, Lin K-Y, Estes A C. Reliability of reinforced concrete girders under corrosion attack. Journal of Structural Engineering 1997; 123(3): 286-297, https://doi.org/10.1061/(ASCE)0733-9445(1997)123:3(286).
16. Ghanooni-Bagha M, Shayanfar M A, Shirzadi-Javid A A, Ziaadiny H. Corrosion-induced reduction in compressive strength of self-compacting concretes containing mineral admixtures. Construction and Building Materials 2016; 113: 221-228, https://doi.org/10.1016/j.conbuildmat.2016.03.046.
17. Gonzales J A, Andrade C, Alonso C, Feliu S. Comparison of rates of general corrosion and maximum pitting penetration on concrete embedded steel reinforcement. Cement and. Concrete Research 1995; 25(2): 257-264, https://doi.org/10.1016/0008-8846(95)00006-2.
18. Li C Q. Reliability based service life prediction of corrosion affected concrete structures. Journal of Structural Engineering 2004; 130(10): 1570-1577, https://doi.org/10.1061/(ASCE)0733-9445(2004)130:10(1570).
19. Li CQ, Zheng J J, Lawanwisut W, Melchers R E. Concrete delamination caused by steel reinforcement corrosion. Journal of Materials in Civil Engineering 2007; 19(7): 591-600, https://doi.org/10.1061/(ASCE)0899-1561(2007)19:7(591).
20. Mirza S A, MacGregor J G. Variability of mechanical properties of reinforcing bars. Journal of the Structural Division 1979; 105 (ASCE 14590 Proceeding): 921-937.
21. Naess A, Leira BJ, Batsevych O. System reliability analysis by enhanced Monte Carlo simulation. Structural Safety 2009; 31(5): 349-355, https://doi.org/10.1016/j.strusafe.2009.02.004.
22. Nogueira C G, Leonel E D. Probabilistic models applied to safety assessment of reinforced concrete structures subjected to chloride ingress. Engineering Failure Analysis 2013; 31: 76-89, https://doi.org/10.1016/j.engfailanal.2013.01.023.
23. Nowak A S, Collins K R. Reliability of structures. CRC Press; 2012.
24. Nowak A S, Szerszen M M, Szeliga E K, Szwed A, Podhorecki P J. Reliability-based calibration for structural concrete. University of Nebraska, UNLCE. 2005:05-3.
25. Östlund L. An estimation of γ-values. Bulletin du Comité Euro-international du Béton 1991(202): 38-97.
26. Papadakis V G, Roumeliotis A P, Fardis M N, Vagenas C G. Mathematical modelling of chloride effect on concrete durability and protection measures. Concrete repair, rehabilitation and protection 1996; 165-174.

27. Pedeferri P. La corrosionenelcalcestruzzo: fenomenologia, prevenzione, diagnosi, rimedi, AICAP, progetto Ulisse, Pubblicemento 2005.
28. Shayanfar M A, Barkhordari M A, Ghanooni-Bagha M. Estimation of Corrosion Occurrence in RC Structure Using Reliability Based PSO Optimization. Periodica Polytechnica. Civil Engineering 2015; 59(4): 531-543, https://doi.org/10.3311/PPci.7588.
29. Shayanfar M A, Barkhordari M A, Ghanooni-Bagha M. Probability calculation of rebars corrosion in reinforced concrete using css algorithms. Journal of Central South University 2015; 22(8): 3141-3150, https://doi.org/10.1007/s11771-015-2851-9.
30. Shayanfar M A, Ghanooni-Bagha M, Jahani E. Reliability theorey of structure. Iust publication Tehran, Iran 2016.
31. Simioni P. Seismic response of reinforced concrete structures affected by reinforcement corrosion (Doctoral dissertation, University of Florence) 2009.
32. Stewart M G, Al-Harthy A. Pitting corrosion and structural reliability of corroding RC structures, experimental data and probabilistic analysis. Reliability Engineering and System Safety 2008; 93(3), 373–382, https://doi.org/10.1016/j.ress.2006.12.013.
33. Stewart M G. Mechanical behaviour of pitting corrosion of flexural and shear reinforcement and its effect on structural reliability of corroding RC beams. Structural Safety 2009; 31(1): 19-30, https://doi.org/10.1016/j.strusafe.2007.12.001.
34. Stewart M G. Spatial variability of pitting corrosion and its influence on structural fragility and reliability of RC beams in flexure. Structural Safety 2004; 26(4); 453–470, https://doi.org/10.1016/j.strusafe.2004.03.002.
35. Stewart M G, Suo Q. Extent of spatially variable corrosion damage as an indicator of strength and time-dependent reliability of RC beams. Engineering Structures 2009; 31(1): 198-207, https://doi.org/10.1016/j.engstruct.2008.08.011.
36. Turnbull A. Review of modelling of pit propagation kinetics. British Corrosion Journal. 2013 Jul 18.
37. Tuutti K. Corrosion of steel in concrete. 1982.
38. Val D V, Melchers R E. Reliability of deteriorating RC slab bridges. Journal of structural engineering 1997; 123(12): 1638-1644, https://doi.org/10.1061/(ASCE)0733-9445(1997)123:12(1638).
39. Vu K A, Stewart M G. Structural reliability of concrete bridges including improved chloride-induced corrosion models. Structural Safety 2000; 22(4): 313-333, https://doi.org/10.1016/S0167-4730(00)00018-7.
40. Vu K, Stewart M G, Mullard J. Corrosion-induced cracking: experimental data and predictive models. ACI Structural Journal 2005; 102(5): 719-726.
41. Zhang X, Wang J, Zhao Y, Tang L, Xing F. Time-dependent probability assessment for chloride induced corrosion of RC structures using the third-moment method. Construction and Building Materials 2015; 76: 232-244, https://doi.org/10.1016/j.conbuildmat.2014.10.039.

Mohammad GHANOONI-BAGHA
Department of civil Engineering, East Tehran Branch
Islamic Azad University
Tehran, Iran

Mohsen Ali SHAYANFAR
The Centre of Excellence for Fundamental Studies in Structural Engineering
Iran University of Science and Technology
P.O.BOX: 16765-163; Narmak, Tehran, Iran

Omid REZA-ZADEH
School of Civil Engineering
Iran University of Science and Technology
P.O. Box 16765-163, Narmak, Tehran, Iran

Masoud ZABIHI-SAMANI
Department of civil Engineering, Parand Branch
Islamic Azad University
Parand, Iran

Emails: ghanoonibagha@iauet.ac.ir, shayanfar@iust.ac.ir,
rezazadeh.omid@yahoo.com, zabihi@piau.ac.ir;

Article citation info:
JODEJKO-PIETRUCZUK A, WERBIŃSKA-WOJCIECHOWSKA S. Development and sensitivity analysis of a technical object inspection model based on the delay-time concept use. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 403–412, http://dx.doi.org/10.17531/ein.2017.3.11.

Anna JODEJKO-PIETRUCZUK
Sylwia WERBIŃSKA-WOJCIECHOWSKA

DEVELOPMENT AND SENSITIVITY ANALYSIS OF A TECHNICAL OBJECT INSPECTION MODEL BASED ON THE DELAY-TIME CONCEPT USE

OPRACOWANIE I ANALIZA WRAŻLIWOŚCI MODELU KONTROLI STANU OBIEKTU TECHNICZNEGO Z WYKORZYSTANIEM KONCEPCJI OPÓŹNIEŃ CZASOWYCH*

In the presented paper, authors focus on the development of mathematical delay-time model for single-unit technical systems (technical objects) liable to costly failure. Failure is taken here to mean a breakdown or catastrophic event, after which the system is unusable until replacement. Implemented maintenance policy is the Block-Inspection Policy that assumes performing inspection actions at regular time intervals of T. In the perfect inspection case the availability and cost models are developed. This gives the possibility for analytical optimization of time between maintenance actions performance T for the infinite operational time of the system. Later, there is examined the compatibility of the developed analytical model with simulation results. The main target is to investigate what is the influence of the given model basic time components on the system availability ratio level and the system long-run expected maintenance costs. The analysis is conducted in the two main steps. The first one regards to analysis of expected number of events (failures, preventive replacements and inspection actions) in a single renewal cycle for the chosen range of time parameters: T and delay time h. In the next step, the availability ratio and long-run maintenance costs dependency on the chosen model's time parameters is under consideration. At the end, the directions for further research work are defined.

***Keywords**: delay time, block-based maintenance, inspection.*

W artykule autorzy skupili się na opracowaniu matematycznego modelu utrzymania obiektów technicznych podlegających kosztownym uszkodzeniom z uwzględnieniem koncepcji opóźnień czasowych. Uszkodzenie w danym przypadku oznacza awarię lub zdarzenie katastrofalne, po którym obiekt jest niezdatny do użytku do momentu wymiany. Wykorzystano politykę blokowej kontroli stanu obiektu, która zakłada, że operacje diagnozy jego stanu są przeprowadzane w regularnych odstępach co T jednostek czasu. Rozpatrzono model kosztowy oraz model gotowości dla przypadku perfekcyjnej diagnozy stanu obiektu. Pozwoliło to na przeprowadzenie analitycznej optymalizacji okresu T między kolejnymi diagnozami stanu obiektu dla nieskończonego horyzontu czasowego. Następnie, zbadano zgodność opracowanego modelu analitycznego z wynikami uzyskanymi w drodze symulacji. Głównym celem było zbadanie wpływu podstawowych parametrów czasowych opracowanego modelu na poziom współczynnika gotowości oraz oczekiwanych kosztów utrzymania badanego obiektu. Analiza została przeprowadzona w dwóch etapach. Pierwszy obejmuje analizę oczekiwanej liczby zdarzeń (uszkodzeń, wymian profilaktycznych oraz operacji kontroli stanu obiektu) dla wybranych zakresów parametrów czasowych: T i opóźnienia czasowego h. W kolejnym kroku zbadano zależność wskaźnika gotowości i oczekiwanych kosztów utrzymania obiektu od wybranych parametrów czasowych modelu. Pracę kończy wskazanie kierunków dalszych prac badawczych.

***Słowa kluczowe**: opóźnienie czasowe, obsługi blokowe, diagnozowanie stanu obiektu.*

1. Introduction

Most of the technical systems are prone to the negative consequences associated with occurrence of unexpected failures due to e.g. degradation processes, environmental conditions, or operator decisions (e.g. [2, 3, 21, 31, 42]). For example, the wear processes are the most common cause of damage (about 70-80% of all the occurred failures) in mechanical devices [22, 41]. Following this, in order to prevent those negative consequences there are introduced some preventive actions related to technical systems maintenance issues.

The tasks connected with appropriate maintenance strategy selection and its parameters optimization are very challenging ones due to the necessity of taking into account a number of requirements connected with processes randomness, limited resources which support operational processes performance, and complexity of analysed technical systems [28, 33]. That is why, over the last 40 years there can be observed an increased interest in problems of technical systems and objects maintenance modelling processes that include e.g. preventive replacement scheduling tasks, inspection policy implementation, or corrective maintenance performance [32, 34]. In particular the pioneering research works dedicated to the maintenance modelling issues especially investigate the performance of single-unit systems (see e.g. works [17, 30, 35]). One of the main research tasks in this area is the optimization of time period between inspection actions performance (e.g. [18]), which is also under authors investigation in this paper.

The authors focus on the modelling of inspection maintenance policy for technical systems/objects based on the technique called a delay time concept (DT). This maintenance approach was developed by Christer et al. (see e.g. [9, 10, 12, 13]). Similarly to Reliability Centred Maintenance strategy (RCM) [20], the delay-time concept is

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

based on a division of technical object's degradation process into two main phases: normal operation (object fully up-stated) and phase of the work, where there can be observed some symptoms of forthcoming failure of components or a system (object partially up-stated). The time period of the second phase is called the delay time, and denoted by h [7]. The known models that base on the implementation of delay-time concept are mainly used for optimizing the time periods between system/object inspection action performances in order to detect some potential signs of future possible failure occurrence. In the case of identification of such symptoms, the technical object is preventively repaired or replaced, which allows for improvement of system dependability characteristics and system total maintenance costs reduction [36]. Recent literature reviews, in which delay-time models are investigated along with other preventive maintenance models, are given in e.g. [16, 23, 26, 30]. Recent state of art works, dedicated strictly to delay-time modelling are given in e.g. [7, 8, 36, 37].

In the area of maintenance modelling of single-unit systems with the use of delay-time approach, one of the first developed works is [13]. Authors in this work investigate a model of equipment-replacement decision process based upon a survey of maintenance managers. The survey research was done in order to investigate how replacement decisions are made in practice. The research study involved the surveys of 19 employees of various organizations. The main focus was to define the criteria which have the biggest influence on maintenance decisions.

Later, in [4], there is analysed the repairable machine that may fail or suffer breakdown many times during the course of its service lifetime, and is inspected for visible faults at intervals. The authors mostly focus on the problem of model parameter estimation with the use of maximum likelihood method and Akaike Information Criterion, providing also a model for imperfect inspections performance.

In work [5] authors develop a two-phase inspection policy model for a single component protection or preparedness system, in which the component arises from a heterogeneous population. The analysed system is required only in an emergency (system on-demand). The modelled two-phase inspection policy bases on the implementation of two inspection frequencies - an anticipated high inspection frequency in early life and low inspection frequency in later life. The optimization process is considered using availability and cost criteria. The policy is investigated in the context of a valve in a natural gas supply network.

Another interesting maintenance model of single-unit system with the use of delay-time approach is presented in work [20]. In the presented paper, author considers a problem of optimal inspection interval determination based on the timeliness measure introduction. The defined measure reflects the expected time between the last inspection and the item failure. Moreover, its minimum defines the optimal inspection interval that can be obtained using Monte Carlo simulation approach.

Moreover, there are also developed in the literature the maintenance models for technical objects that introduce a novel assumption - the so-called postponed replacement. An example here may be the work [33], where authors propose a new delay time model that allows replacement of object with detected defect to be postponed for an additional time period. The key motivation is to achieve better utilization of the system's useful life, and reduce replacement costs by providing a sufficient time window to prepare maintenance resources. The model is analysed for the cases when the delay time is a deterministic and random variable.

Delay time models of single-unit systems reliability are developed in [1, 6, 11, 19, 43]. In [11], author presents simple reliability model of a single component subject to one type of inspect able defect, when inspections are assumed to be perfect. Author focuses on the analysis of the influence of the length of time between inspection actions performance on system reliability characteristics. Then the model is generalized by introducing the possibility of the appearance of the different types of defects (occurring independently in the system). This problem is later investigated in [6], where author simplifies the developed model and provides some numerical examples. The same model given in [11] is provided in [1], when delay time and time to failure densities are exponentially distributed. Later, in [19], authors extend a delay time model with periodic testing process by developing availability function in accordance with the length of time between inspection actions performance. Authors analyse three different variants of the model: the case with no mechanical possibility to detect a defect, zero delay time occurrence, or no inspection actions performance during maintenance processes of the investigated technical object. In the next work, [43], authors investigate the model to evaluate the reliability and optimise the inspection schedule for a multi-defect component. There is also considered the situation of non-constant inspection intervals.

The developed delay-time based maintenance models for single-unit systems under condition based maintenance are presented e.g. in works [14, 24, 25]. In [14], authors analyse an inspection modelling problem for a production plant under condition monitoring. In the analysed article a condition monitoring test checks the state of wear of a component and records a (0,1) signal depending upon all being well or that wear is below a critical level with the possibility of imminent failure. Following this, in the optimization model the decision variables are the critical warning level and frequency of monitoring inspections. Later, in works [24, 25], authors develop a method for determining the discrete time points of inspection for a deteriorating single-unit system under condition-based maintenance. The delay-time model is here utilized to describe the transition of the system's states. Moreover, two types of probabilities with respect to inspections are considered – a failed-dangerous probability of type I error, and a failed-safe probability of type II error.

In the known literature, there can be also found some models dedicated to single-component systems' risk analysis of maintenance activities performance (see e.g. [38]), semi-Markov processes implementation (see e.g. [15]), or safety inspection processes optimization (see e.g. [39]). The example of delay time model for a single unit case is given in [40], where authors propose a method to find cost-optimized maintenance of an elevator.

The previously presented in the literature delay-time-based maintenance cost and availability models were mainly limited to the analysis of the first inspection period (the time between the moment of the "as good as new" technical object performance beginning to the moment of its first inspection action performance). In practice, this would mean the definition of the time of the first inspection action performance and the lack of guidelines for further maintenance procedures performance in the case of not detecting of symptoms of forthcoming failure (e.g. DT models presented in [27]). The use of Block inspection policy in real-life operating systems involves periodic inspection actions performance. This makes that it is necessary to determine the time period between inspection actions performance that is the best from the point of view of maintenance costs minimizing in the long run.

Therefore, the article focuses on the development of a mathematical model for a technical object maintenance, which extends the approach used so far and allows analysing the long-term operation time period of a single-unit system. The model gives the possibility to find a constant time period between the inspection actions performance that is optimal due to minimal maintenance costs or maximal availability criterion satisfaction. Depending on the inspection action results, there are taken necessary actions (maintenance and/or further operation). The next inspection is performed after the next time period T regardless of the earlier inspection results. This significantly simplifies the management process of technical object maintenance.

Moreover, there is presented an analysis of the sensitivity of the developed model to the selected parameters that gives the possibility to indicate certain relationships, which can facilitate the search for the optimal period between inspection actions performance due to the both considered criteria.

As a result, in the next Section there are presented the modelling problem description and main assumptions. This gives the possibility for the investigation of the mathematical maintenance model for technical unit that base on Block Inspection policy implementation. The optimization criteria include the long-run expected maintenance costs and availability ratio for the given technical object. Later, there is examined the compatibility of the developed analytical model with simulation results. Then there is conducted the sensitivity analysis and later there is analysed the possibility of using the presented maintenance model to optimize the technical object's inspection period. The work ends with a summary.

2. Problem description and main assumptions

2.1. Problem description

In the presented paper, there is analysed a technical object with time delay, subject to periodic inspection maintenance. The conducted research study focuses on objects that are unrepairable or repairable but their corrective maintenance performance is costly ineffective. At the same time, it is assumed that the technical object has a single failure mode, and symptoms of forthcoming failure are identifiable during inspection actions performance.

Based on the relevant scientific literature, here we analyse a single-unit system subject to periodic inspection. An implemented maintenance policy is the Block Inspection policy. This maintenance policy assumes that the performed inspections are carried out in order to check the working status of the object and take place at regular time intervals of T. The purpose of such inspection action is to establish if the technical object is still operable. Due to the simplicity of this type of maintenance policy, it is still often recommended by manufacturers. An example of this type of maintenance policy may regard to the periodic inspections of hydraulic pressure in e.g. wheel loaders. Such break pressure checking is performed in every 1000 machine working hours.

Moreover, the inspections of technical object are assumed to be perfect. In the presented article the inspections are assumed to be perfect. As a result, object defect, which occurs till the moment of inspection, will be identified and replaced within the inspection period. At the same time, it is assumed that if these symptoms are diagnosed, the object will be replaced by a new one. Moreover, failure is observed immediately and the object is replaced at a given cost and downtime. The main decision variable in the model is the inspection interval T, which is called the maintenance cycle. The objective of the model is to define and estimate the availability ratio of the object and the long-term expected maintenance costs for a single renew period for the case of the random nature of the exploitation process (random moments of defects and failure occurrence). For the modelling purpose, we propose the following additional assumptions.

2.2. Assumptions

In the presented work, the general assumptions we make to characterize the operation and maintenance of the non-repairable technical unit are listed below:

1) the technical object is a three state object where, over its service life, it can be:
   - operating – able to fulfil the operational tasks and functions (later called as full up-state),
   - partially operating (denoting the existence of a defect in the object that is defined as the occurrence of the symptoms of potential failure - in the further part of the study called as incomplete or partial up-state), or
   - down for necessary repair (after the technical object failure occurrence, defined as the loss of the ability of the object to perform the required functions),
2) the object is renewed (replaced by a new one that has identical reliability characteristic) at either a failure repair or at a repair done at an inspection if a defect is identified,
3) failures of the object are identified immediately,
4) replacements are made as soon as possible and have a constant time. The time required for the object replacement is determined by variables d_f and d_r in Subsection 3.1.,
5) the object inspection action lasts a constant time (determined by a variable d_{in} described in Subsection 3.1.) and begins the new inspection cycle T for the analysed object (presented graphically in the Fig. 2).

3. Inspection model for a single-unit system based on the delay-time concept implementation

Here authors investigate a case of a technical object performing under Block Inspection policy principles. In general, the analysed object may be in one of the three maintenance operations: failure repair without additional inspection action performance, planned inspection performance with replacement of an element if a defect is detected, or planned inspection action performance without replacement of an element (defect not detected, object assumed to be fully up-stated). Thus, if the inspection action reveals a failure or a defective state of the object, a repair or replacement is initiated immediately otherwise it remains in service.

For the needs of this research, the period between the moment of the new technical object operation beginning to the moment of its replacement is called the "renewal cycle", while the time elapsing between two successive inspections (or inspection and successive system failure) is called the "inspection cycle". Figure 1 illustrates the definition of the initial and delay times in the case of one component maintenance. Figures 2 and 3 present the renewal and defect renewal cycle for a technical objects.

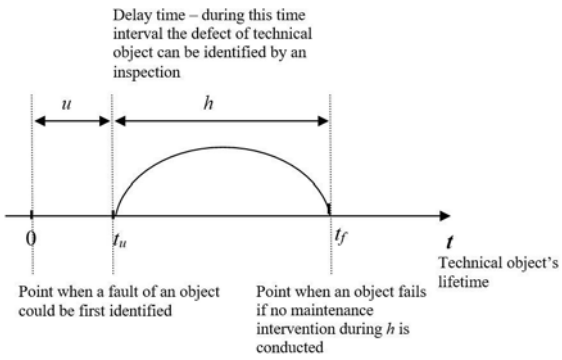


Fig. 1. The initial and delay times of a single-unit technical system failure process. Source: Own contribution based on [8]

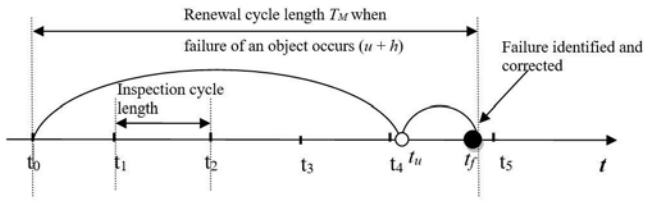


Fig. 2. Failure renewal cycle length of an analysed single-unit technical system

For the presented model assumptions, there can be estimated the availability ratio and the long-term maintenance costs for a given re-

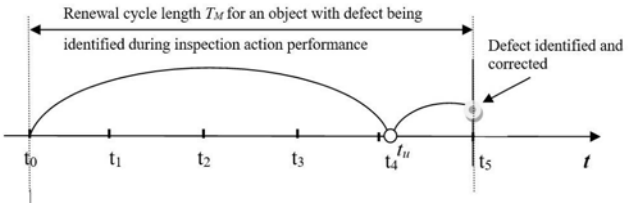


Fig. 3. Defect renewal cycle length of an analysed single-unit technical system

newal cycle (composed of inspection cycles of length T), giving a foundation to look for the best value of the T period. In order to determine these characteristics, the modelling efforts will be subjected to analysis of one renewal cycle of a technical object. In practice, this approach makes it possible to estimate the costs and the availability of the object in an infinite horizon time period, because the determined probabilities and expected values will be the same for subsequent renewal cycles. This is due to the fact that after the end of the analysed/modelled renewal cycle, the object is replaced (preventively or correctively) with a new one with identical reliability characteristics.

The availability and cost results need to estimate the probabilities of the object being in the possible reliability states in a single renewal period, like object uptime, object partially uptime (there are signs of forthcoming failure) and system downtime (failure) in the long term (when $t \to \infty$). Determination of these probabilities, taking into account the fact of the object inspection performance, gives the possibility in the next step to estimate the expected values of its maintenance costs and its reliability characteristics.

3.1. Notation

In order to develop an analytical delay-time model we shall adopt the following notation:

h	the delay time of the defect, denoting the period between the moment of appearance of the first symptoms of the potential failure and the moment of the object failure occurrence
u	initial time of the defect
t_u	the moment of the first symptoms of the potential failure occurrence
t_f	the random moment of failure occurrence in the object
t_i	the moment of the ith inspection action performance, where $i = 1,2,3,\ldots, \infty$
$t_i \to \infty$	the moment of the ith inspection action performance for $i \to \infty$ (infinite horizon time)
T_i	time of ith inspection cycle performance (between two consecutive inspection actions performance)
$F(t)$	Cumulative distribution function of an object's time to failure t
$f_h(h)$	Probability density function of the delay time h
$F_h(h)$	Cumulative distribution function of the delay time h
$g(u)$	Probability density function of the initial time u, which elapses from the beginning of operation by an "as good as new" object until the moment of first symptoms of failure occurrence
$G(u)$	Cumulative distribution function of the initial time u, which elapses from the beginning of operation by an "as good as new" object until the moment of first symptoms of failure occurrence
$R(t)$	Reliability function of an object being operated by time t
$P_I(t_i)$	probability that during the ith inspection action performance, an object is still up stated
$L_I(ti \to \infty)$	the expected number of inspection actions performance in a single renewal cycle of an object, during which it will be up stated
$P_R(ti \to \infty)$	the probability that an object will be replaced during a single renewal cycle
$P_F(ti \to \infty)$	the probability of corrective replacement of an object in a single renewal cycle
$L_R(ti \to \infty)$	the expected number of preventive replacements performance in a single renewal cycle
$L_F(ti \to \infty)$	the expected number of an object failures occurrence in a single renewal cycle
$T_M(t_1)$	the expected lifetime of an object (up state and partially up state time periods) in the first inspection cycle
$T_M(t_i \to \infty)$	the expected lifetime of an object (up state and partially up state time periods) from the moment of its operation beginning to the moment of its replacement (the length of a renewal cycle according to the Fig. 2 and 3)
$R_h(t_1,t_2)$	the probability that in the time period between time moments t_1, t_2 the symptoms of forthcoming failure will occur, but an object will not fail up till the time moment t_2
$T_I(t_i \to \infty)$	the expected time of inspections carried out in an object during a single renew period
d_{in}	single inspection action performance time
$T_R(t_i \to \infty)$	the expected preventive replacement time of an object during a single renew period
d_r	object's defect replacement time
$T_F(t_i \to \infty)$	the expected failure replacement time of an object during a single renew period
d_f	object's corrective replacement time (after failure occurrence)
$A(t_i \to \infty)$	the availability ratio of an object for a single renew period
$C_I(t_i \to \infty)$	the expected cost of inspection action performance per a single renewal cycle
c_{in}	unit inspection cost of an object
$C_R(t_i \to \infty)$	the expected cost of preventive replacement performance in a single renewal cycle
c_r	unit replacement cost of an object
$C_F(t_i \to \infty)$	the expected cost of corrective replacement performance in a single renewal cycle
c_f	unit consequence costs incurred in the case of object failure
$C(t_i \to \infty)$	the long-term expected maintenance costs per unit time

3.2. Probabilities of a system uptime, partially uptime and failure in a single renewal cycle

Assuming that the failure of the object is always preceded by detectable symptoms, there is a possibility to define the Cumulative distribution function of time to failure, $F(t)$, as the convolution of u and h such that $u + h \le t$ [27]:

$$F(t) = \int_{u=0}^{t} g(u) F_h(t-u)\, du \tag{1}$$

And the reliability: $R(x) = 1 - F(x)$

The probability $P_I(t_i)$ that during the ith inspection action performance of the technical object (which started working as a "new" one) it will still be in the up state, can be defined as:

$$P_I(t_i) = \left(1 - G(t_i)\right) = 1 - \int_0^{t_i} g(u)\, du, \text{ where: } t_i = T, 2T, 3T, \ldots, iT \tag{2}$$

Based on this, it is possible to determine the expected number of inspection actions performance in a single renewal cycle, during which the object will be up-stated $L_I(t_i \to \infty)$:

$$L_I\left(t_i \to \infty\right) = \sum_{i=1}^{\infty}\left(P_I\left(t_i\right)\right) = \sum_{i=1}^{\infty}\left(1 - G\left(t_i\right)\right) = \sum_{i=1}^{\infty}\left[1 - \int_0^{t_i} g\left(u\right)du\right] \tag{3}$$

The second possible situation that can occur during the object inspection action performance regards to the necessity of its replacement due to existing signals of the potential forthcoming failure. In order to illustrate the possibility of preventive replacement occurrence during the modelled (single) renewal cycle, the Figure 4 shows alternative possibilities of the occurrence of a sequence of events that cause the preventive replacement maintenance during the first three inspection cycles, carried out within one renewal cycle.

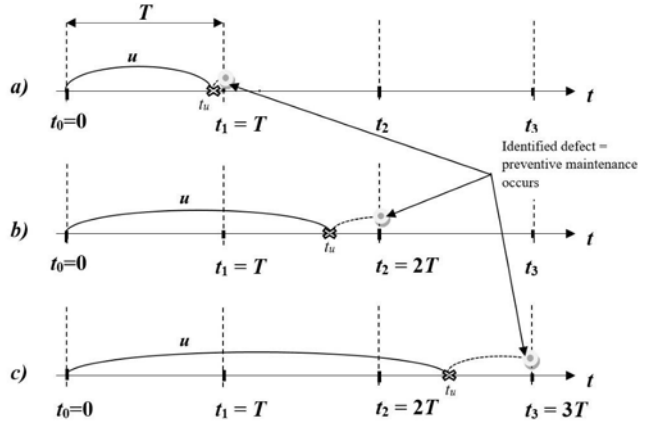


Fig. 4. Scheme for the renewal cycle length, when the symptom of potential system's failure (defect) is identified during inspection action performance in the first (a), second (b) and third (c) inspection cycle

Figure 4 shows the examples of the versions of the random time period durations: up to the occurrence of defect (u) and the delay time (h), which would result in the preventive replacement performance at the time T (Fig. 4a), at the time $2T$ (Fig. 4b), or at the $3T$ (Fig. 4c). The probability of each of the variants occurrence (a, b, c) may be determined for one inspection cycle. For example, for the second inspection cycle (Fig. 4b), the probability that symptoms appear during his lifetime, and the object will be preventively replaced at the time $2T$ can be described as:

$$P_R\left(t_i = 2T\right) = \int_{t_1}^{t_2} g\left(u\right)\left(1 - F_h\left(t_2 - u\right)\right)du \tag{4}$$

Based on this, there is possible to estimate the probability of the object preventive replacement in the infinite renewal cycle. The probability $P_R(t_i \to \infty)$ that an item should be replaced preventively can be described as the sum of the probabilities that during subsequent inspection actions performance there will be observed symptoms of future possible failure occurrence:

$$P_R\left(t_i \to \infty\right) = \sum_{i=1}^{\infty}\left(R_h\left(t_{i-1}, t_i\right)\right) = \sum_{i=1}^{\infty}\left[\int_{t_{i-1}}^{t_i} g\left(u\right)\left(1 - F_h\left(t_i - u\right)\right)du\right] \tag{5}$$

Due to the fact that preventive replacement is univocal to the end of the analysed renewal cycle, the series given by the equation (5) converges to an asymptotic value describing the probability of preventive replacement performance during one inspection cycle.

The estimation of cost and reliability characteristics of the analysed object also needs to determine the possibility of its failure. The probability that the object will fail during the inspection cycle depends on whether there will be visible symptoms of forthcoming failure and failure occurs before the next inspection action performance. As in the case of preventive replacement, the probability of the technical object failure is calculated for successive periods of maintenance, carried out within one renewal cycle, which ends with the corrective replacement action performance. Examples of possibilities of the failure occurrence during $t = (0, t_3)$ are shown in the Figure 5.

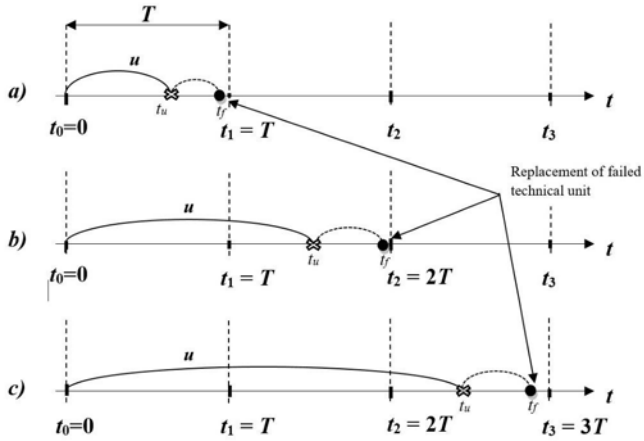


Fig. 5. Scheme for the renewal cycle length, when system's failure occurs in the first (a), second (b) and third (c) inspection cycle

The probability of corrective replacement of the object in a single renewal cycle may be estimated by:

$$P_F\left(t_i \to \infty\right) = \sum_{i=1}^{\infty}\left[\int_{t_{i-1}}^{t_i} g\left(u\right)\left(F_h\left(t_i - u\right)\right)du\right] \tag{6}$$

The variables $P_R(t_i \to \infty)$ and $P_F(t_i \to \infty)$ represent the probabilities of preventive and corrective replacement necessity in the object as well as they express the expected number of possible preventive and corrective replacements in a single renewal cycle of the object. This is connected with the adopted assumption that during the whole renewal cycle there is possible only one replacement of the object, which determines the moment of renewal cycle end:

$$L_R\left(t_i \to \infty\right) = 1 * P_R\left(t_i \to \infty\right) \tag{7}$$

$$L_F\left(t_i \to \infty\right) = 1 * P_F\left(t_i \to \infty\right) \tag{8}$$

3.3. Availability ratio for a system

In the curse of the first inspection cycle the expected lifetime of the object may be expressed as:

$$T_M\left(t_1\right) = \int_0^{t_1}\int_0^{x} g\left(u\right) f_h\left(x - u\right)dudx + t_1 \cdot \int_{u=0}^{t_1} g\left(u\right)\left(1 - F_h\left(t_1 - u\right)\right)du \tag{9}$$

The first component of the sum in the formulae (9) expresses possible moments of object failures, while the second part of the expression takes into consideration the probability that the object defect is

observed in the time of the first inspection performance (then there is a necessity to replace it preventively).

The total expected length of the object up time in the curse of a single renewal cycle may be estimated as the sum of expected lifetimes resulting from the sequence of inspection periods, when $t_i \to \infty$:

$$T_M(t_1) = \int_0^{t_1} x \int_0^{x} g(u) f_h(x-u)\,du\,dx + t_1 \cdot \int_{u=0}^{t_1} g(u)\left(1 - F_h(t_1 - u)\right)du \quad (10)$$

In order to estimate the expected length of the object renewal time period there should be determined the expected times of all maintenance operations being performed in a single renewal cycle.

The next inspection (ith) may be taken depending on the probability of the two facts: that a previous inspection has not revealed defect occurrence and at the end of (i–1)th inspection period the object is still reliable. Following this, the expected time of inspections carried out in the object during a single renew period takes into account the probabilities $P_R(t_i \to \infty)$ and $P_F(t_i \to \infty)$. Thus, when one wants to estimate the expected time of inspection actions performance in a single renew cycle, the formulae (11) may be applied:

$$T_I(t_i \to \infty) = d_{in} \cdot \left(L_I(t_i \to \infty) + L_R(t_i \to \infty)\right) \quad (11)$$

The estimation of the expected time of preventive replacement of the technical object in the course of a single renew period should base on the probability of defect occurrence at the end of an inspection cycle and lack of this defect at the moment of the previous inspection performance:

$$T_R(t_i \to \infty) = d_r \cdot L_R(t_i \to \infty) \quad (12)$$

Moreover, there should be estimated the expected failure time in a single renew cycle. The time depends on the failure probability in a given inspection period, assuming that at the moment of a previous inspection performance any defect has not been detected:

$$T_F(t_i \to \infty) = d_f \cdot L_F(t_i \to \infty) \quad (13)$$

The calculated functions (10)-(13) allow estimating the availability ratio for the cycle $A(t_i \to \infty)$, according to the well-known formulae:

$$\mathrm{A}(t_i \to \infty) = \frac{\mathrm{T_M}(t_i \to \infty)}{\mathrm{T_M}(t_i \to \infty) + \mathrm{T_I}(t_i \to \infty) + \mathrm{T_F}(t_i \to \infty) + \mathrm{T_R}(t_i \to \infty)} \quad (14)$$

3.4. The long-run expected maintenance costs of a system

The estimation of expected maintenance cost results of the object needs the definition of the basic cost components connected with inspection performance, preventive replacements and corrective replacement. For the single renewal cycle, these basic cost components may be estimated similarly as time characteristics:

$$C_I(t_i \to \infty) = c_{in} \cdot \left(L_I(t_i \to \infty) + L_R(t_i \to \infty)\right) \quad (15)$$

$$C_R(t_i \to \infty) = \mathrm{c}_r \cdot L_R(t_i \to \infty) \quad (16)$$

$$C_F(t_i \to \infty) = \mathrm{c}_f \cdot L_F(t_i \to \infty) \quad (17)$$

Thus, the long-term expected maintenance costs per unit time $C(t_i \to \infty)$ may be defined according to the well know formulae:

$$C(t_i \to \infty) = \frac{C_I(t_i \to \infty) + C_R(t_i \to \infty) + C_F(t_i \to \infty)}{T_M(t_i \to \infty)} \quad (18)$$

3.5. The convergence of analytical and simulation results of the presented maintenance model

In order to present the convergence of the foregoing models and simulation ones, presented in the literature (see e.g. [18]), in this Subsection authors focus on the comparison of the obtained results for general case, taking into account the technical object exemplary parameters.

The Figure 6 shows the total expected length of the object up time in the curse of a single renew cycle (TM) calculated analytically by equation (10), and using a Monte Carlo simulation. As it can be seen, the results are characterized by a very large convergence, and the maximum relative error for the presented case is 4% of the "simulation cycle length". Other obtained results are characterized by a similar convergence, regardless of the given parameters of the object.

In addition to comparing the results, the Figure 6 shows that the two test parameters T and h have a significant impact on the length of the object lifetime. Shorter time intervals between inspection actions performance (shorter T) mean that the occurred defects in the object are more often identified during inspection. This also affects the length of the object operating time by preventive replacement performance (lower values of TM).

Change of the parameter h in the analysed case means the change of time u, after which a defect occurs in the item. It was assumed, that the tested objects are characterized by the same expected operational time ($h + u = 100$ = constant) and there are changed only the lengths of each "sub-period" u and h. Increasing value of h means that the signal of impending failure can be observed earlier and for a longer period than in the case of decreasing h. In the examined case, the maximum value of h ($h = 90$) indicates that the defect signal is observable in the object for about 9 of 10 parts of its time to failure. In the Figure 6, it can be seen that in this case frequent object inspections reduce several times the potential object operational time in the system through preventive replacement performance. Such an undesirable effect can be seen in the Figure 7, which shows the availability ratio for the same random variable of T and h. Frequent inspections in combination with the long-time of defect presence, cause a decrease in the object's availability for even adopted simplifying assumption of the equality of all time components ($d_i = d_r = d_f = 1$).

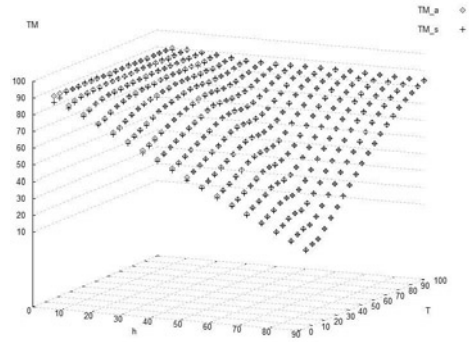


Fig. 6. The analytical (TM_a) and simulation-based (TM_s) expected lifetime of an element for various values of expected delay time (h) and various lengths of times between inspections (T)

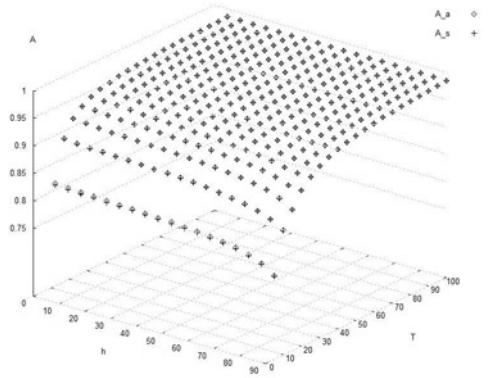


Fig. 7. The analytical (A_a) and simulation-based (A_s) system availability ratio for various values of expected delay time (h) and various lengths of times between inspections (T)

These results illustrate how huge is the need for determining the time between inspection actions performance of the object for its main characteristics, as system reliability features, time components, or cost elements (Fig. 7). For this reason, in the next Section, there is presented a brief sensitivity analysis of the results obtained on the basis of the developed maintenance models.

4. Sensitivity analysis of the developed availability and cost models for a single-unit system with delay-time

In this Section, there is presented the short sensitivity analysis of the models developed above. The objective of the given analysis was to investigate what is the influence of given model basic time components on the object availability ratio level and maintenance costs in the infinite time horizon.

The sensitivity analysis of the developed cost and availability modes is begun by analysing the expected number of performed maintenance actions (failures (equation (8), preventive replacements (equation (7), and inspection actions (equation (3)), for various values of expected delay time (h) and various lengths of times between inspections (T) – Figures 8-10.

As we can see in the Figures 8-10, the expected number of performed inspection actions (Fig. 8) is inversely proportional to the rate of failure of the object due to the fact that only in the event of the object failure the renewal cycle does not end by inspection action performance. When the time of a defect occurrence is short ($h \approx$ 5-15), frequent inspections ($T \approx$ 5-10) cause that there is a necessity to perform them many times in a single renewal cycle (Fig. 8). On the other

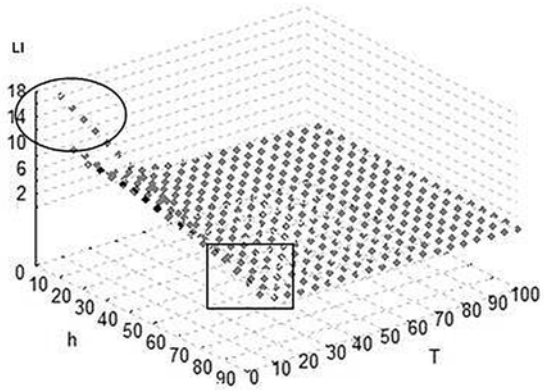


Fig. 8. The expected number of performed inspection actions (regardless of the result) in the single renew cycle for various values of expected delay time (h) and various lengths of times between inspections (T)

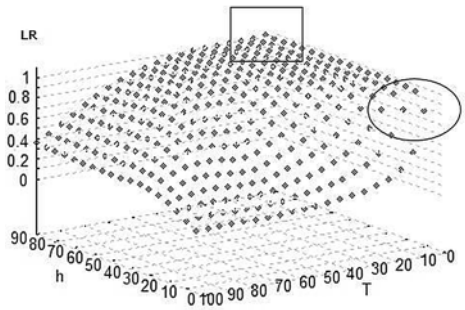


Fig. 9. The expected number of system's element preventive replacements in the single renew cycle for various values of expected delay time (h) and various lengths of times between inspections (T)

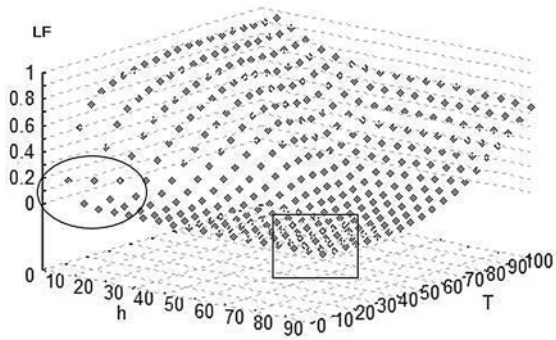


Fig. 10. The expected number of system's element corrective replacements in the single renew cycle for various values of expected delay time (h) and various lengths of times between inspections (T)

hand, a large number of inspections results in higher effectiveness of inspection activities, as it significantly reduces the likelihood of object failure occurrence (Fig. 10). These cases are marked in the graphs by ellipses. However, the same time period between inspections T, used in the case of the objects with a long delay time ($h \approx$ 50-90), does not significantly alter the probability of object failure, which remains low even in fewer inspections performance ($T \approx$ 20-30), but increases significantly the number of object preventive replacements (Fig. 9), which is indicated in the graphs by a rectangle.

As we can see in the Figures, for the tested values of T and h the expected numbers of inspections actions performance, preventive replacements and object failures have a large area of variation, which means that they are sensitive to both analysed parameters. While we are mostly not able to control the length of the object delay time (h), it is possible to control the length of time between inspections actions performance T in a system. However, the user has to answer the question of what criteria should and can be taken into account when optimizing time period T. Thus, the selected results on cost and reliability, achieved through the test object are presented in the Figures 11-14.

Graphs, presented in the Figures 11-14, show the object maintenance costs and its availability for the two chosen cases, where the unit inspection cost of an object and inspection time (c_i, d_i) are relatively important compared with the preventive and corrective replacement times and costs (Figs. 11, 12), or where these parameters are of a little importance compared to the inspection and failure costs (Figs. 13, 14). There can be seen in all of these Figures, that for any length of delay time h there is a value of the period T, which gives better modelling results than for other lengths of T (in Figures these values are marked by arc for the tested range of parameter h). A thorough analysis of the results confirms that the best T period from the point of view of minimizing the object maintenance costs is close to the half of the delay time h length. The relationship is similar to the Shannon rule that regards to continuous-time signals sampling. In the case of the objects inspection action performance, if the test procedure will

be carried out at least twice in the expected delay time period, the chance to avoid costly failure to the object is the largest, and the costs are the smallest.

In the case of the availability ratio maximization, the best times between inspections (T) stronger depend on the time units associated with each event in the object than it is observable for the cost minimization. When the single inspection action performance time (d_i) is relatively small compared to the time of failure (d_r) (Fig. 14), the solution that maximizes the availability of the object is achieved with more frequent inspections than in solutions optimizing the costs of the object maintenance (Fig. 13) for the same ratio of costs and time units.

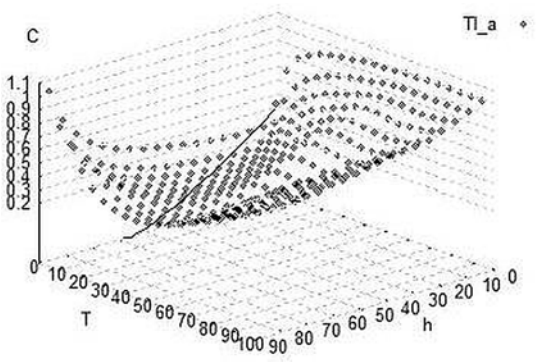


Fig. 11. The system expected maintenance costs for various values of expected delay time (h) and various lengths of times between inspections (T), assuming: $c_i = 1$, $c_r = 10$, $c_f = 100$

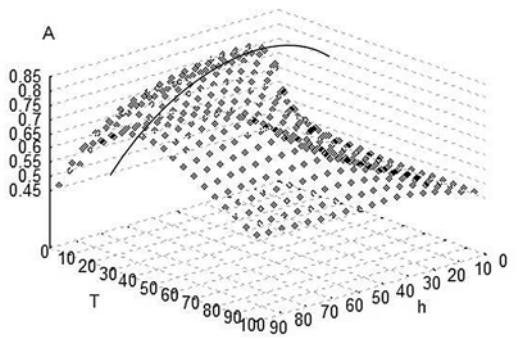


Fig. 12. The system availability ratio for various values of expected delay time (h) and various lengths of times between inspections (T), assuming: $d_i = 1$, $d_r = 10$, $d_f = 100$

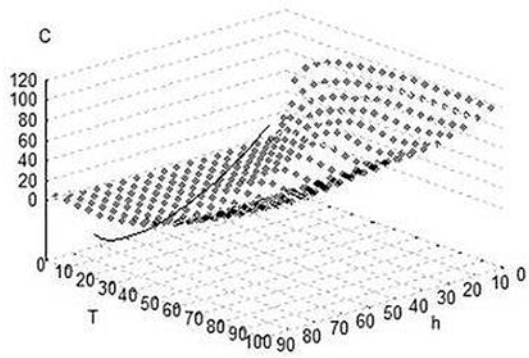


Fig. 13. The system expected maintenance costs for various values of expected delay time (h) and various lengths of times between inspections (T), assuming: $c_i = 1$, $c_r = 100$, $c_f = 1000$

The presented examples show that the best time between inspections of the object must be determined individually for each change of reliability, cost, or time parameters and taking into account the chosen optimization criterion. Only in some cases, it may happen that the same period T will give the best results in terms of both the analysed criteria.

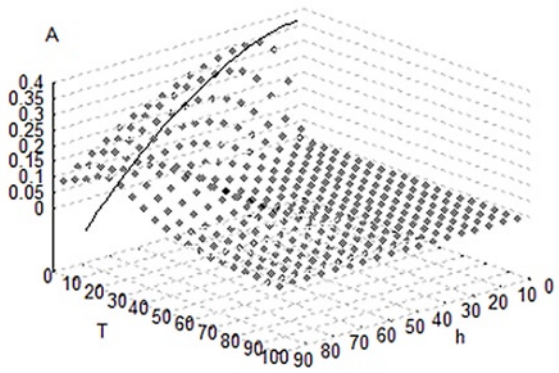


Fig. 14. The system availability ratio for various values of expected delay time (h) and various lengths of times between inspections (T), assuming: $d_i = 1$, $d_r = 100$, $d_f = 1000$

5. Summary

The paper presents analytical models describing the availability and maintenance costs of the technical object based on delay-time approach use. These models may be used to optimize the length of time between successive inspection actions performance T for the given cost and reliability characteristics.

Presented in the literature known analytical maintenance models that base on the delay-time analysis implementation usually are limited to an analysis of the time period, from the moment of system's new element ("as good as new") launching to the moment of the first inspection action performance. The models proposed in this article provide the basis for analytical identification of the optimal length of time T in the infinitely long time horizon of the technical object operation and maintenance. The paper also presents a brief sensitivity analysis of the obtained models for the selected object parameters, which led to the conclusion that the best lengths of time T, taking into account the maximization of the object availability, are not always equal to the best lengths of this time period being obtained from the point of view of obtaining the minimal maintenance costs of the given object. At the same time, the conducted research shows that for the various parameters of the object, there is only one global minimum of the cost function or maximum of its availability. Thus, this gives a possibility to search for the optimal time period T with the use of widely known optimization methods, like anti-gradient methods of optimizing functions of one variable. This is a significant convenience compared to optimization solutions that based on simulation models.

The practical application of the presented model may be difficult, especially because of the assumption of perfect inspection case investigation. This is often seen in practice that the inspection action performance may not give the right answer if and what defect occurs in the system. This may be caused by various factors such as the skill of the inspector, the accuracy of the inspection tools, etc. Moreover, the maintenance decision may have a multidimensional nature and depend on the problems connected with definition if and what kind of signal of forthcoming failure may occur. Such a situation may lead to different decisions made by managers for the same maintenance situation. Thus, this issue will be the topic of authors' future papers.

References

1. Attia A. F. Estimation of the reliability function using the delay-time models. Microelectronics Reliability 1997; 37(2): 323-327, https://doi.org/10.1016/S0026-2714(96)00012-1.
2. Babiarz B. An introduction to the assessment of reliability of the heat supply systems. International Journal of Pressure Vessels and Piping 2006; 83(4): 230-235, https://doi.org/10.1016/j.ijpvp.2006.02.002.
3. Bajda A., Wrażeń M., Laskowski D. Diagnostics the quality of data transfer in the management of crisis situation. Electrical Review 2011; 87(9A): 72-78.
4. Baker R. D., Wang W. Estimating the delay-time distribution of faults in repairable machinery from failure data. IMA Journal of Mathematics Applied in Business & Industry 1992; 3: 259-281.
5. Cavalcante C. A. V., Scarf P. A., de Almeida A. T. A study of a two-phase inspection policy for a preparedness system with a defective state and heterogeneous lifetime. Reliability Engineering and System Safety 2011; 96: 627-635, https://doi.org/10.1016/j.ress.2010.12.004.
6. Cerone P. On a simplified delay time model of reliability of equipment subject to inspection monitoring. Journal of the Operational Research Society 1991; 42(6): 505-511, https://doi.org/10.1057/jors.1991.98.
7. Christer A. H. A Review of Delay Time Analysis for Modelling Plant Maintenance. in: Stochastic Models in Reliability and Maintenance, Osaki S. (ed.), Springer, 2002, https://doi.org/10.1007/978-3-540-24808-8_4.
8. Christer A H. Developments in delay time analysis for modelling plant maintenance. Journal of the Operational Research Society 1999; 50: 1120-1137, https://doi.org/10.1057/palgrave.jors.2600837.
9. Christer A. H. Delay-time model of reliability of equipment subject to inspection monitoring. Journal of the Operational Research Society 1987; 38(4): 329-334, https://doi.org/10.1057/jors.1987.54.
10. Christer A. H. Modelling inspection policies for building maintenance. Journal of the Operational Research Society 1982; 33: 723-732, https://doi.org/10.1057/jors.1982.161.
11. Christer A. H., Waller W. M. A Descriptive model of capital plant replacement. Journal of the Operational Research Society 1987; 8(6): 473-477, https://doi.org/10.2307/2582760.
12. Christer A. H., Waller W. M. Reducing production downtime using delay-time analysis. Journal of the Operational Research Society 1984; 35(6): 499-512, https://doi.org/10.1057/jors.1984.103.
13. Christer A. H., Waller W. M. Delay Time Models of Industrial Inspection Maintenance Problems. Journal of the Operational Research Society 1984; 35(5): 401-406, https://doi.org/10.1057/jors.1984.80.
14. Christer A. H., Wang W. A model of condition monitoring of a production plant. International Journal of Production Research 1992; 30(9): 2199-2211, https://doi.org/10.1080/00207549208948145.
15. Christer A.H., Wang W., Choi K., Van der Duyn Schouten F. A. The robustness of the semi-Markov and delay time single-component inspection models to the Markov assumption. IMA Journal of Management Mathematics 2001; 12: 75-88, https://doi.org/10.1093/imaman/12.1.75.
16. Dekker R., Scarf P. A. On the impact of optimisation models in maintenance decision making: the state of the art. Reliability Engineering and Safety 1998; 60: 111-119, https://doi.org/10.1016/S0951-8320(98)83004-4.
17. Frostig E. Comparison of maintenance policies with monotone failure rate distributions. Applied Stochastic Models in Business and Industry 2003; 19: 51-65, https://doi.org/10.1002/asmb.485.
18. Jodejko-Pietruczuk A., Nowakowski T., Werbińska-Wojciechowska S. Time between inspections optimization for technical object with time delay. Journal of Polish Safety and Reliability Association, Summer Safety and Reliability Seminars 2013; 4(1): 35-41.
19. Jia X. Christer A. H. A periodic testing model for a preparedness system with a defective state. IMA Journal of Management Mathematics 2002; 13: 39-49, https://doi.org/10.1093/imaman/13.1.39.
20. Jiang R. A timeliness-based optimal inspection interval associated with the delay time model. Proc. of Prognostic and System Health Management Conference PHM - 2012, Beijing, 2012, https://doi.org/10.1109/phm.2012.6228965.
21. Kierzkowski A., Kisiel T. Functional readiness of the check-in desk system at an airport. Theory and engineering of complex systems and dependability. Proc. of the Tenth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX, June 29 - July 3, 2015, Brunów, Poland. Springer, 2015;. 223-233, https://doi.org/10.1007/978-3-319-19216-1_21.
22. Laskowski D., Bylak M. Efficient diagnostics encoding mechanism for wireless networks. Electrical Review 2013; 89(9): 133-138.
23. Mazzuchi T. A., van Noortwijk J. M., Kallen M. J. Maintenance optimization. Technical Report, TR-2007-9, 2007.
24. Okumura S. An Inspection Policy for Deteriorating Processes Using Delay-Time Concept. International Transactions in Operational Research 1997; 4(5-6): 365-375, https://doi.org/10.1111/j.1475-3995.1997.tb00092.x.
25. Okumura S., Jardine A. K. S., Yamashina H. An inspection policy for a deteriorating single-unit system characterized by a delay-time model. International Journal of Production Research 1996; 34(9): 2441-2460, https://doi.org/10.1080/00207549608905037.
26. Ozekici S. (ed.). Reliability and Maintenance of Complex Systems. NATO ASI Series, Series F: Computer and Systems Sciences, 154. Springer, 1996.
27. Redmond D. F. Delay Time Analysis in Maintenance. PhD thesis. Salford: University of Salford; 1997.
28. Restel F.J. The Markov reliability and safety model of the railway transportation system. Safety and reliability: methodology and applications: proceedings of the European Safety and Reliability Conference, ESREL 2014, 14-18 September, 2015, Wrocław, Poland. CRC Press/ Balkema: 303-311.
29. Tang Y., Jing J. J., Yang Y., Xie C. Parameter estimation of a delay time model of wearing parts based on objective data. Mathematical Problems in Engineering 2015, https://doi.org/10.1155/2015/419280.
30. Valdez-Flores C., Feldman R. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. Naval Research Logistics 1989; 36: 419-446, https://doi.org/10.1002/1520-6750(198908)36:4<419::AID-NAV3220360407>3.0.CO; 2-5.
31. Valis D., Koucky M., Žak L. On approaches for non-direct determination of system deterioration. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14(1): 33-41.
32. Valis D., Vintr Z. Vehicle Maintenance Process Optimisation Using Life Cycle Costs Data and Reliability-Centred Maintenance. Proceedings of the First International Conference on Maintenance Engineering. Beijing: Science Press, 2006.

33. Van Oosterom C. D., Elwany A. H., Celebi D., van Houlum G. J. Optimal policies for a delay time model with postponed replacement. European Journal of Operational Research 2014; 232: 186-197, https://doi.org/10.1016/j.ejor.2013.06.038.
34. Vintr Z., Valis D. A Tool for Decision Making in k-out-of-n System Maintenance. Applied Mechanics and Materials 2012; 110-116: 5257-5264.
35. Wang H. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 2002; 139: 469-489, https://doi.org/10.1016/S0377-2217(01)00197-7.
36. Wang W. An overview of the recent advances in delay-time-based maintenance modelling. Reliability Engineering and System Safety 2012; 106: 165-178, https://doi.org/10.1016/j.ress.2012.04.004.
37. Wang W. Delay time modelling. In: Kobbacy K. A. H., Prabhakar Murphy D. N. (eds.). Complex system maintenance handbook. Springer-Verlag London Limited, 2008: 345-373, https://doi.org/10.1007/978-1-84800-011-7_14.
38. Wang W. A delay time based approach for risk analysis of maintenance activities. Journal of the Safety and Reliability Society 2003; 23(1):103-113, https://doi.org/10.1080/09617353.2002.11690753.
39. Wang W., Christer A. H. A modelling procedure to optimize component safety inspection over a finite time horizon. International Quality and Reliability Engineering 1997; 13: 217-224. https://doi.org/10.1002/(SICI)1099-1638(199707)13:4<217::AID-QRE107>3.0.CO;2-P
40. Yamashina H., Otani S. Cost-optimized maintenance of the elevator - single unit case. Journal of Quality in Maintenance Engineering 2001; 7(1): 49-70. https://doi.org/10.1108/13552510110386946
41. Zajac M., Swieboda J. Process hazard analysis of the selected process in intermodal transport. International Conference on Military Technologies (ICMT), 19-21 May 2015; 1-7. https://doi.org/10.1109/miltechs.2015.7153698
42. Zieja M., Wazny M., Stepien S. Distribution determination of time of exceeding permissible condition as used to determine lifetimes of selected aeronautical devices/systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(1): 57-64. https://doi.org/10.17531/ein.2016.1.8
43. Zhao J., Chan A. H. C., Roberts C., Madelin K. B. Reliability evaluation and optimisation of imperfect inspections for a component with multi-defects. Reliability Engineering and System Safety 2007; 92: 65-73. https://doi.org/10.1016/j.ress.2005.11.003

Anna JODEJKO-PIETRUCZUK
Sylwia WERBIŃSKA-WOJCIECHOWSKA
Faculty of Mechanical Engineering
Department of Operation and Maintenance of Logistic Systems, Transportation Systems and Hydraulic Systems
Wroclaw University of Science and Technology
ul. Wybrzeze Wyspianskiego 27, 50-370, Wroclaw, Poland

E-mails: anna.jodejko@pwr.edu.pl, sylwia.werbinska@pwr.edu.pl

Article citation info:
PAVLOVIĆ P, MAKAJIĆ-NIKOLIĆ D, VUJOŠEVIĆ M. A new approach for determining the most important system components and the budget-constrained system reliability improvement. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 413–419, http://dx.doi.org/10.17531/ein.2017.3.12.

Petar PAVLOVIĆ
Dragana MAKAJIĆ-NIKOLIĆ
Mirko VUJOŠEVIĆ

A NEW APPROACH FOR DETERMINING THE MOST IMPORTANT SYSTEM COMPONENTS AND THE BUDGET-CONSTRAINED SYSTEM RELIABILITY IMPROVEMENT

NOWE PODEJŚCIE DO WYZNACZANIA NAJWAŻNIEJSZYCH ELEMENTÓW SYSTEMU ORAZ POPRAWY NIEZAWODNOŚCI SYSTEMU W WARUNKACH OGRANICZONEGO BUDŻETU

Importance measures are used for indexing system components due to their impact on the system's overall reliability. In order to identify the specific number of the most critical components, first-ranked components are singled out as the most important ones. However, importance measures consider only the influence of individual components and they are not applicable to combinations or groups of components. This common feature of importance measures is referred to in literature as one of still open issues. This paper proposes a new approach for determining the most important system components, where a whole set of components are determined simultaneously taking into account their interdependence. In systems with a large number of interdependent components, the number of the most important components which should be prevented is often limited due to the available budget. Using pre-known minimal cut sets, a mathematical model based on the Budgeted Maximum Coverage Problem is proposed. By its optimization, the simultaneous determination of all of the most important components whose total expenses do not exceed the limited overall budget is achieved. The new approach was tested by a series of experiments conducted over a set of test examples. The results of the experiments were compared with the results obtained using two commonly used cost importance measures.

Keywords: *reliability, importance measures, critical components, budgeted maximum coverage problem, optimization.*

W złożonych systemach, w których koszty poprawy niezawodności poszczególnych elementów są znane, często ogranicza się budżet przeznaczony na podnoszenie ogólnej niezawodności systemu. W takich przypadkach konieczna jest maksymalizacja niezawodności systemu przy jednoczesnym utrzymaniu kosztów na poziomie minimum. Powszechnie znane metody rozwiązywania powyższego problemu opierają się na wyznaczaniu ważności kosztów, co wymaga określenia rang elementów składowych systemu, a w dalszej kolejności wyodrębnienia pewnej liczby najważniejszych elementów pierwszorzędnej rangi. W niniejszej pracy zaproponowano nowe podejście do określania najważniejszych komponentów systemu w oparciu o problem maksymalnego pokrycia w granicach budżetu (budgeted maximum coverage problem); podejście wdrażano z wykorzystaniem wcześniej znanych minimalnych przekrojów niezdatności. Optymalizacja proponowanego modelu matematycznego, pozwoliła na jednoczesne wyznaczenie wszystkich najważniejszych elementów, dla których łączne wydatki na utrzymanie ruchu nie przekraczały całkowitego ograniczonego budżetu. Nowe podejście zostało przebadane w serii eksperymentów przeprowadzonych na zbiorze przykładów testowych, za które posłużyły wzorcowe drzewa błędów. Wyniki badań porównano z wynikami uzyskanymi za pomocą dwóch miar ważności kosztów – miary ważności opartej na kosztach oraz miary ważności opartej na opłacalności. W większości przypadków, proponowany model dawał lepsze wyniki niż pomiary ważności kosztów.

Słowa kluczowe: *niezawodność, miara ważności, elementy krytyczne, problem maksymalnego pokrycia zbioru w granicach budżetu, optymalizacja.*

1. Introduction

Improvement of the reliability of a complex system, besides providing redundancy of components, is achieved by improving the reliability of the individual components that have the greatest impact on the overall reliability of the system. In this paper, we consider a complex system as a system with multiple interactions between many different components [15] in which the behavior of the system cannot be easily and directly induced from the behavior of its components [3]. Within the context of reliability, main characteristics of a complex system are: large number of components, cut sets, path sets or components' states dependencies [7]. Analysis of the impact of the components' reliability on the overall system reliability is usually done by using importance measures. The concept of importance measures, originally introduced by Birnbaum [2] and Vesely [23], has been significantly changed and improved over time. Structural and probabilistic importance measures are used most often, but a small number of importance measures that take into account the costs of improving the reliability of the individual components have also been developed over the last few years [12, 24]. A common feature of the majority of importance measures is that component's importance to the system functioning and failure is determined based on the calculated value of a chosen importance measure [27]. Therefore, the general procedure

of their application is: firstly, the numerical value for the importance of each individual component is calculated using the chosen importance measure formulae; secondly, components are ranked considering previously calculated values. In the case a desired number k of the most important components need to be determined, it would be the k first-ranked components since the higher ranked components are considered the most critical. There are some importance measures available in the literature that do not fully share the mentioned common feature. Joint Reliability Importance (JRI) analyses a pair of components [1], Differential Importance Measure (DIM) is calculated as a sum of the individual importance measures [4], and enxtention of DIM called DIM^{II} combines JRI and DMI [29]. However, according to Zio [28], the fact that the majority importance measures rank only individual components so that they are not directly applicable to combinations or groups of components, is still one of the open issues on importance measures.

The subject of this paper is the formulation and validation of a new importance measure which, unlike the existing approaches, besides considering costs of individual components takes into account their mutual impact on the overall system reliability. The new measure is formulated as an optimization problem, more precisely as the Budgeted Maximum Coverage Problem (BMCP). BMCP is presented by Khuller as the variation of Maximum Coverage Problem [13], but in literature, it is also recognized as a Maximum Coverage with Knapsack Constraints [8]. The approach found its application in graphs [5, 21], location problems [19] and healthcare networks monitoring [6]. This paper could indicate the first usage of BMCP in reliability since the authors of this paper have not been able to find any previous applications in this area. There are examples in the literature where information provided by importance measures is used in the formulation of an optimization problem in order to obtain a system design in which all components have similar importance values [30]. However, according to our knowledge, optimization approach so far has not been used to determine the importance of the components. Using the new proposed optimization approach all of the most critical components are determined simultaneously, and the sum of their individual costs fits the available budget. The approach is experimentally tested and the obtained results verifying the approach are presented and explained.

The proposed approach is based upon minimal cut sets (MCSs) of coherent fault trees, so the terminology of the fault tree analysis is used. Primary events, which represent the system components' failures, are elements of MCSs. Cut sets and MCSs are defined as follows:

Definition 1 [9]: Cut set is a set of events that together cause the top undesired event to occur.

Definition 2 [9]: Minimal cut set (MCS) is a cut set reduced to a minimum number of events that cause the top undesired event to occur.

MCSs are most commonly generated based on a fault tree of the observed system [16]. Formally, a fault tree is a directed and connected acyclic graph G=(N,A), where N is the set of nodes and A is the set of arcs (Fig. 1).

The top event of the fault tree (root node TE), represents a system failure. The failure modes of the system components are primary events (components' failures) represented by leaves e1-e6. Besides intermediate nodes G1-G4, a given fault tree has four logic operators that model the cause-effect relationships between components' failures - two OR operators are assigned to G1 and G2 while two AND operators are assigned to nodes G3 and G4. For the fault tree shown in Figure 1, MCSs are: e1e4, e1e5e6, e2e3e4 and e2e3e5e6. Each of these combinations of primary events could cause the system failure. The qualitative analysis (in the process of Boolean reduction of a set of equations) identifies the minimal cut sets which are the combinations of the smallest number of primary events and if they

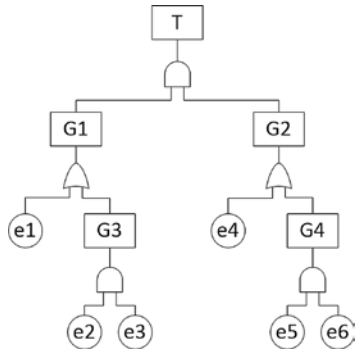


Fig. 1. Fault tree example

occur simultaneously, they may lead to a top event. The top event can be expressed as [26]

$$TE = \bigcup_{i=1}^{n} MCS_i = \bigcup_{i=1}^{n} \left(\bigcap_{j=1}^{m} P_j \right)_i \tag{1}$$

where TE is the top event, MCS_i is the i-th minimal cut set, n is the total number of minimal cut sets, P_j is the j-th primary event and m is the total number of primary events. The quantitative analysis represents a calculation of the top event probability. Considering the assumption that the primary events are mutually independent, the top event probability of occurrence Q_{TE} may be approximated as [26]

$$Q_{TE} = \sum_{i=1}^{n} Q_{MCS_i} = \sum_{i=1}^{n} \left(\prod_{j=1}^{m} Q_{P_j} \right)_i \tag{2}$$

where Q_{MCS_i} is the probability of occurrence of the minimal cut set i and Q_{P_j} is the probability of occurrence of the primary event j. The proposed approach assumes that all MCSs of a given fault tree are already determined.

The paper is organized as follows: Section 2 deals with importance measures that consider the cost of components used for the validation of a proposed approach. The problem of simultaneous determination of the most critical system components constrained with alimited budget is formulated as BMCP in Section 3. A mathematical model of a formulated optimization problem and solving methods are presented in Section 4. Section 5 gives the comparison of the results obtained by the proposed model with the results obtained using the importance measures from Section 2, tested on a group of benchmark fault trees. Concluding remarks and discussion are presented in Section 6.

2. Importance measures that consider costs of components

The importance measures indicate the system components that have the greatest impact on the reliability of the system [10]. Overviews of some of the most used importance measures are available in many books and papers [14, 20, 22].

The two importance measures used in this paper to compare results with our proposed approach are Cost-based component importance (CBCI) and Cost-effective importance measure (CEIM).

CBCI is introduced as the extension of Birnbaum importance [25]. In our paper CBCI of component i is defined as:

$$I_i^{CBCI}(t) = -\frac{\partial C_i(t)}{\partial R_i} \tag{3}$$

where $\partial C_i(t)$ and ∂R_i represent the increase of system cost and system reliability caused by reliability improvement of i-th component, respectively. CBCI can be interpreted as follows: when $I_i^{CBCI}(t)$ is large, a small change in the reliability of component i will result in a comparatively large change in the total cost of maintaining the entire system during the time interval $(0,t)$.

CEIM combines the concept of reliability importance measure and the total costs of failure, and it is defined as:

$$I_i^{CEIM}(t) = \frac{I_i^{GI}(t)}{C_{f,i}} \tag{4}$$

where I_i^{GI} is the general importance (GI) of component i at time t, and $C_{f,i}$ is a cost factor for i-th component. GI of component i is calculated as:

$$I_i^{GI}(t) = \frac{\Delta g_i(Q(t))}{g(Q(t))} \tag{5}$$

where $\Delta g_i(Q(t))$ represents the change in system probability due to change in i -th component probability. The cost factor for i-th component is calculated as the ratio of sum of the expected costs of failure for all components and the expected cost of reliability improvement for i-th component:

$$\Delta g_i(Q(t)) \tag{6}$$

where $E(C_i)$ is the expected cost of reliability improvement for i-th component. A component which gives a maximum benefit at minimum costs will be termed as the most cost-effective component and will possess the highest rank in priority list [12].

3. Problem formulation

In this paper, we observe the improvement of components reliability as a decrease of component's failure probability to a very low or negligible value, close to zero. This assumption appears in most importance measures and represents a decreased risk level with the component optimized or assumed to be perfectly reliable, i.e. components reliability equal to 1 [20].

If we observe the fault tree given at Figure 1, we can assign costs for their six primary components, for example: c1=5, c2=6, c3=7, c4=8, c5=7 and c6=6 cost units (c.u.). Also we can assume some limited budget, for example B=15 c.u., available for improvement of the most critical system components. In that case, both CBCI and CEIM can be calculated using equations (3) and (4) in order to find which components should be improved. After calculation of values for each component, we are able to make rankings presented in the table below:

Both measures basically gave the same output. They selected components e1 and e4 as the most critical, and their total costs of improvement of 13 c.u. fit the available budget. If we calculate the overall system reliability improvement, we get apercentage of 99.17, which is a good result. This improvement is obtained by comparing the initial system reliability with the reliability of the system in which selected components (e1 and e4) are assumed to be perfectly reliable.

Table 1. Rankings of components' importance using CBCI and CEIM

Rank	Component ranking by CBCI	Calculated CBCI value	Component ranking by CEIM	Calculated CEIM value
1.	e1	40.58	e4	0.19
2.	e4	64.94	e1	0.12
3.	e6	175.93	e3	0.02
4.	e2	175.93	e5	0.02
5.	e3	205.28	e2	0.01
6.	e5	205.28	e6	0.01

Observing the list of MCSs from the example given in figure 1, it can be seen that primary events e1 and e2 are present in every MCS. Therefore, if those two components were selected for improvement of their reliability i.e. the probability of their failure is set to zero, then regarding equation (2) the probability of occurrence of TE would also equal zero. Thus, by selecting components e1 and e2 the percentage of 100 % of system reliability improvement can be achieved. Components e1 and e2 fits limited budget even better, as their total costs are only 11 c.u.

In addition, if the available budget is increased, the results of CBCI and CEIM could be used just to add next lower ranked component to the set of the previously selected most critical components with a higher ranking. The approach proposed in this paper in case of increasing the available budget is able to find completely different set K of k critical components which achieve higher overall system reliability improvement, while still fitting the budget. The traditional way of calculating importance measures by making independent calculations for each individual component and linear expansion of set K of k critical components in case the available budget is increased, are issues that are overcome by the approach proposed in this paper.

The new approach of determining the most important system components relies on minimal cut sets defined in the introduction chapter. The basic implication of definitions 1 and 2 is - if any of components/events which are the element of an MCS can be prevented to fail, then MCS stops to be the cause of possible system failure and is considered eliminated or "covered". The probability of realization of MCS is then used to diminish the overall system failure probability, i.e. system reliability is improved at the same rates. The goal is to select components which should be improved so that their non-failure would maximize the reduction of the overall system failure probability.

Starting presumptions are:

- Minimal cut sets of the observed system are known
- Probabilities for failure of systems' components are given (or their order of magnitude)
- Cost of improvement is known for each individual component
- Total budget available for system reliability improvement is given and limited

The proposed approach can be formulated as an optimization problem described as follows: allocate the available budget to ensure non-failure of components which eliminate the most probable MCSs i.e. which maximize the reduction of the system failure probability.

4. Mathematical model and solving methods

Mathematical model of described optimization problem is made using the following notation:

- S - set of primary events;
- $i \in S$ - primary event which represents the failure of i-th system component;
- m - the number of MCSs;
- w_j – weight, i.e. probability for the realization of j-th minimal cut set (calculated by multiplying failure probabilities of its components)
- c_i – cost for preventing i-th component from failure, i.e. improving it for non-failure
- B – available budget for system reliability improvement
- a_{ij} - binary variable defined as:

$$a_{ij} = \begin{cases} 1, \text{ if } i\text{-th component is an element of } j\text{-th minimal cut-set} \\ 0, \text{ othervise} \end{cases}$$

– x_i - binary variable called a disabling indicator associated with the primary event i, $i \in S$, defined as:

$$x_i = \begin{cases} 1, & \text{if the primary event } i \text{ is disabled (componetnt functioning)} \\ 0, & \text{othervise (component failure)} \end{cases}$$

– y_j - binary variable called elimination indicator associated with the MCS j, j=1,…,m, defined as:

$$y_i = \begin{cases} 1, \text{if } j\text{-th minimal cut-setis disabled (covered)} \\ 0, \text{othervise} \end{cases}$$

The mathematical model of allocating limited budget to prevent failure of those components which eliminate the most probable minimal cut sets, is formulated as follows:

BMCP (budgeted maximum coverage model):

$$max\ f(x,y) = \sum_{j=1}^{m} w_j y_j \tag{7}$$

s.t.

$$\sum_{i \in S} a_{ij} x_i \geq y_j\ ,\ j = 1,\ldots,m \tag{8}$$

$$\sum_{i \in S} C_i x_i \leq B \tag{9}$$

$$x_i \in \{0,1\}, i \in S$$

$$y_j \in \{0,1\}, j = 1,\ldots,m$$

Objective function maximizes the overall probability of eliminated MCSs which implicate the increase of system reliability. The first constraint ensures elimination of all of the MCSs which contain any failure prevented components. The second constraint is related to the available budget. Considering the first constraint and the maximization of the objective function, the binary requirement for a variable can be relaxed with . The formulated mathematical model corresponds to the Budgeted Maximum Coverage Problem [13] as well as a Maximum Coverage with Knapsack Constraints [8].

Mathematical model BMCP was solved in two ways. Firstly, it was solved exactly using GLPK solver [11], and then the possibility of failure equal to zero was assigned to all selected events that should be prevented; using equation (2) a new system reliability and the percentage of reliability improvement is calculated. Secondly, it was solved using a heuristic developed for the observed problem, called GreedyPlus. The proposed heuristic uses a variation of the greedy algorithm, slightly improved to be able to recognize and remove components which cover redundant MCSs. Thus the overall system reliability could be improved by selecting some new components whose failures should be prevented. For each event, weights of all MCSs that contain specific event are summarized and based on that, events are sorted in a non- decreasing array. Events are selected, starting from the first one of that series until the budget is exhausted. Algorithm also picks the cheapest components in the case of a tie. After each selection of event, all MCSs containing the selected event are eliminated, i.e. set to 0, and the remaining events after each step form a new array based on the weights of the remaining uncovered MCSs. Adaptation of greedy algorithm is reflected in the fact that the reduction procedure after each step has to check whether any of the already selected events is unnecessary, i.e. whether it covers exactly the same MCSs as some of the already selected events. In this case, the solution is reduced, i.e. such an event is discarded. The budget is increased for the amount of cost of discarded event so a new event needs to be selected. Since the BMCP is NP-hard [13], it can be difficult to find its optimal solution for large fault treesin areasonable time. For this purpose, GreedyPlus could be used as the starting point for future development of some special heuristics that would deal with the mentioned problem. Algorithm of GreedyPlus and pseudo-codes for the used procedures are given below. Notation used in algorithm description is:

W_i – the sum of the weights of MCSs that contain i-th event, $i \in S$,
L – non-decreasing array consisting of non-redundant calculated values W_i, $i \in K$.
c_i – cost for i-th component being prevented from failure
B – available budget for system improvement

Algorithm 1. GreedyPlus

```
greedy()
count single coverings and form array L
reduction()
```

Procedure 1.greedy (MCS[m][l], c[i], B)

```
end_signal←0
repeat
    i←1
    repeat
        j←1
        repeat
            if MCS[i][l]=j then
                W[i] ← W[j]+poss[j]
            if MCS[i][j]=0 or B=0 then
                end_signal←1
            if W[i]>max then
                max←W[i]
                X[F] ←i
                B←B-c[i]
                MCS[i][l] ←0
        until j=n
    until i=m
until end_signal=1
```

Procedure 2. reduction (X[F], Lmin, appearanceNo[F])

```
singleCover[]←0
i←1
repeat
    if X[i]=appearanceNo[i] then
        singleCover[i]← singleCover[i]+1
until i= F
i←1
repeat
    if singleCover[i]=0 then
        singleCover[i]← singleCover[i]+1
        B←B+c[i]
        X[i] ←Ø
until i= n
greedy (MCS[m][l], c[i], B)
```

Both solutions of BMCP model, the optimal solution obtained by GLPK and the solution obtained using GreedyPlus, were compared with solutions obtained by CBCI and CEIM.

5. Experimental results

The proposed new approach is first ilustrated on the fault tree of train rear-end collision accident and then on a group of benchmark fault trees.

The fault tree of train rear-end collision accident, retrieved from [17], consists of 35 primary and 17 intermediate events. The observed fault tree has 24000 MCSs which means that top event called "train rear-end collision" has 24000 failure modes. The rank, i.e. the number of primary events of 18000 MCSs is nine while the rank of the remaining 6000 MCSs is eight.

In order to apply the proposed approach, the following assumptions were introduced: the probabilities of all primary events are 0.01; the available budget for system reliability improvement is 60 cost units; and the costs for preventing components from failure are between 5 and 30 cost units. In addition, since events 1 and 2 from [17] appear in all MCSs, those two events are immediately identified as the most important and further analysis was made for the remaining 33 events.

Presented mathematical model (7-9) was first solved exactly using GLPK solver [10] and the percentage of reliability improvement is calculated according to the optimal solution. Then, reliability improvement percenteges obtained by CBCI and CEIM importance measures are calculated in the following steps:

step 1: Using equations (3) and (4), different ranks of primary events were obtained.

step 2: The possibility of failure eaqual to zero was assigned to the top ranked primary events whose total costs do not exceed the available budget.

step 3: A new system reliability was calculated.

step 4: Percentages of improvement of the system relialibility were calculated.

Optimization using GLPK and steps 1 to 4 were repeated for ten instances of randomly generated costs for preventing components from failure. The obtained results are presented in Table 2. Notation used in Table 2 is: BS = budget spent (percentage), NC = number of components, and RI = reliability improvement (percentage).

The results in Table 2 show that the proposed approach outperforms both importance measures in all ten instances. In eight instances, reliability improvement of 100% is achieved which means that obtained events cover all MCSs. In the case of CEIM and CBCI, such reliability improvement is obtained only in one instance. This is directly caused by the fact that those measures calculate only individual contribution to system reliability and do not consider the impact of groups of components. On the other hand, based on the proposed approach, a set of the most important components often contains components whose individual impact is not among the best ranked. For example, in the first instance, the set of the most critical events consists of events: 7, 8, and 9, according to the optimal solution of (7-9). The event 8 is second ranked by CEIM, event 9 is sixst ranked by CBCI, while the event 7 is not among the first ranked by any of them. However, their combination provides the highest reliability improvement. Similar observations apply for all other instances.

Mathematical model (7-9) and the proposed algorithm are afterwards verified by experiments conducted used as test examples [18]. Characteristics of those benchmark fault trees (BFT) are given in Table 3:

Table 3. Benchmark fault trees

BFT name	E	BE	MCS	R
das9201	204	122	14217	2-7
das9202	85	49	27778	1-11
baobab2	72	32	4805	2-6
baobab3	187	80	24386	2-11

The column labeled with E gives the total number of events contained in fault tree, while BE column gives the number of primary events (components' failure) i.e. events that form minimal cut sets. Total number of minimal cut-sets in the fault tree is given in the column MCS. Column R gives ranges of ranks for minimal cut-sets (e g. a tree named das9201 has the smallest MCS's rank of 2, while the highest rank of MCSs in that BFT is 7, and so on).

Ten random instances of costs of improvement for primary events/components are generated for each BFT given in Table 3. For each instance four available budgets were tested, so it was made 160 experiments in total. Each experiment gave four outputs for each of tested methods (CBCI, CEIM, GLPK and GreedyPlus). Average values of the experimental results are presented in Table 4.

Notation used in Table 4 is given bellow:

ABS = average budget spent (cost units)

ANC = average number of components (*avg* for 10 instances, each inst. outputed *int* value)

ARI = average reliability improvement (percentage)

B = available budget (cost units)

Observing the obtained results, it can be seen that GLPK and GreedyPlus gave better, or at least the same quality outputs as CBCI and CEIM, i.e ARI value in columns labeled GLPK and GreedyPlus are generally much larger than values in CBCI and CEIM columns. Due to the way of calculation, CBCI tends to select a greater number of cheaper components while CEIM prefers a smaller number of expensive ones, but neither of those two traditionally calculated importance measures managed to give better results than our newly proposed approach applied in GLPK and GreedyPlus.

Also, it should be noted that it is possible to achieve the same system reliability improvement with the same amount of budget but with aselection of a different number of critical components. If so, all outputs of such kind are actually multiple solutions and have equal quality, but it simply looks wiser to choose a solution with fewer critical components and less time needed for their prevention.

Table 2. Results of ten instances for the fault tree of train rear-end collision accident

	GLPK			CEIM			CBCI		
instances	NC	RI	BS	NC	RI	BS	NC	RI	BS
1	3	100	78.33	3	60	98.33	8	85.60	100
2	4	100	90	2	46.67	100	7	74.54	95
3	4	100	100	3	73.33	100	8	90.40	95
4	4	100	95	3	66.67	100	8	82.28	100
5	4	100	100	3	75	100	7	82.28	90
6	3	100	98.33	3	100	98.33	7	84.41	98.33
7	8	98.95	100	3	60	100	8	85.60	86.67
8	4	100	98.33	2	40	93.33	7	88	100
9	5	84.85	98.33	2	50	100	6	76.62	100
10	4	100	98.33	3	73.33	98.33	7	83.04	88.33
average	4.30	98.38	95.67	2.70	64.50	98.83	7.30	83.28	95.33

Table 4. Experimental results given as average values of results obtained on ten instances

		CBCI			CEIM			GLPK			GreedyPlus		
		ABS	ANC	ARI	ABS	ANC	ARI	ABS	ANC	ARI	ABS	ANC	ARI
das9201	B=26	23.4	4.4	18.79	23.9	1.1	14.33	24.8	2.9	30.73	24.5	2.8	30.65
	B=53	49.3	8.7	27.77	52.2	2.7	38.63	52.6	4.8	51.89	51.5	4.6	51.69
	B=106	102.8	16	44.89	104.5	6	75.72	104.7	7.4	82.41	104.2	7.1	82.33
	B=160	157.2	22.5	52.95	159.1	8.5	92.20	154.4	8.9	96.64	146.1	8.6	96.64
das9202	B=10	6.1	1.1	6.02	7.8	1	14.16	7.8	1	14.16	7.6	1	14.16
	B=21	18.1	3	9.65	20.5	1.7	44.14	20	2	46.09	19.9	1.8	46.03
	B=42	38.2	5.8	12.31	41.3	2.6	87.29	41.3	2.8	87.61	41.3	2.7	87.60
	B=64	59.7	8.2	16.69	62.6	3.6	92.84	62.8	4.1	96.15	62.6	4	96.08
baobab2	B=7	5.3	1	13.66	5.8	1	14.82	5.8	1	14.82	5.5	1	14.82
	B=14	12	2	28.07	12.4	1.1	22.01	12.5	1.9	29.09	12.2	1.7	27.18
	B=28	26.3	4.1	43.59	26.6	1.7	32.83	27	3.3	51.64	26.7	2.8	50.75
	B=42	38.8	5.6	56.45	40.7	2.3	45.70	40.9	4.2	65.86	40.6	4.1	65.26
baobab3	B=17	15.8	3	13.81	15.1	1.1	16.76	15.1	2	23.95	15	1.5	20.85
	B=35	31.7	5.4	25.68	34.7	2	25.83	34.3	3.6	40.62	33.7	2.5	35.17
	B=70	66.3	10.2	37.24	68.3	3.1	41.93	68.9	5.7	62.48	68.7	4.3	57.67
	B=105	101.4	14.3	46.07	103.3	4.4	55.32	103.6	6.9	79.09	103.5	6.3	76.81

Comparing GLPK outputs with the outputs of GreedyPlus, it can be observed that the improved greedy algorithm reached an average reliability improvement obtained by the exact algorithm only in three cases (das9201 with B=160, das9202 with B=10 and baobab2 with B=7), while all other results are very close to optimal. These three cases achieved smaller ABS value due to the way in which the improved greedy algorithm treats multiple solutions – it always picks a cheaper one, while GLPK just takes care about the goal function i.e. the overall system improvement.

6. Conclusion

The problem of determining the set of the most critical system components in case of a limited budget has been solved by a new approach which formulates a problem as a Budgeted Maximum Coverage Problem. The approach is successfully tested and verified by a series of experiments over benchmark fault trees. It is concluded that in cases where MCSs are pre-known, it is more convenient to apply this new approach rather than importance measures as CBCI and CEIM.

The proposed approach could be limited by the impossibility of obtaining exact solutions in a reasonable time for the problems of large dimensions. Therefore, in a further work the presented improved greedy algorithm could be incorporated into specially designed heuristics, possibly based on Variable neighborhood search (VNS).

References

1. Armstrong M. Joint reliability importance of elements. IEEE Transactions on Reliability 1995; 44(3): 408-412, https://doi.org/10.1109/24.406574.
2. Birnbaum Z. On the importance of different components in a multicomponent system. In P. Krishnaiah (Ed.), Multivariate Analysus-II. New York: Academic Press 1969.
3. Bar-Yam Y. Dynamics of complex systems (Vol. 213). Reading, MA: Addison-Wesley 1997.
4. Borgonovo E, Apostolakis G. A new importance measure for risk-informed decision making. Reliability Engineering and System Safety 2001; 72: 193-212 ,https://doi.org/10.1016/S0951-8320(00)00108-3.
5. Caskurlu B, Mkrtchyan V, Perekh O, Subramani K. On Pratial Vertex Cover and Budgeted Maximum Coverage Problems in Bipartite Graphs. Theoretical Computer Science: 8th IFIP TC 1/WG 2.2 International Conference, TCS 2014, Rome, Italy: Springer 2014; 13-25.
6. Curtis DE, Pemmaraju SV, Polgreen P. Budgeted Maximum Coverage with Overlapping Costs: Monitoring the Emerging Infections Network. 2010 Proceedings of the Twelfth Workshop on Algorithm Engineering and Experiments (ALENEX). Society for Industrial and Applied Mathematics 2010.
7. Der Kiureghian A, Song J. Multi-scale reliability analysis and updating of complex systems by use of linear programming. Reliability Engineering and System Safety 2008; 93(2): 288-297, https://doi.org/10.1016/j.ress.2006.10.022.
8. Du D, Ko K, Hu X. Design and analysis of appromaxition algorithms. Springer Optimization and Its Applications 2012, https://doi.org/10.1007/978-1-4614-1701-9.
9. Ericson II C A. Hazard Analysis technique for System Safety. New Jersey: John Wiley & Sons 2015.
10. Espitrity J, Coit D, Prakash U. Component criticalty importance measures for the power industry. Electric Power Systems Research 2007; 407-420, https://doi.org/10.1016/j.epsr.2006.04.003.
11. GLPK (GNU Linear Programming Kit) From: www.gnu.org/software/glpk.

12. Gupta S, Bachttacharya J, Barabady J, Kumar U. Cost-effective importance measure: A new approach for resource prioritization in a production plant. International Journal of Quality & Realibility Management 2013; 30 (4): 379-386, https://doi.org/10.1108/02656711311308376.
13. Khuller S, Moss A, Naor J. The budgeted maximum coverage problem. Information Processing Letters 1999; 70 (1): 39-45, https://doi.org/10.1016/S0020-0190(99)00031-9.
14. Kuo W, Zhu X. Importance measures in reliability, risk and optimization. Chichester: John Whiley & Sons 2012, https://doi.org/10.1002/9781118314593.
15. Ladyman J, Lambert J, Wiesner K. What is a complex system? European Journal for Philosophy of Science 2013; 3(1): 33-67, https://doi.org/10.1007/s13194-012-0056-8.
17. Li Y F, Mi J, Huang H Z, Zhu S P, Xiao N. Fault tree analysis of train rear-end collision accident considering common cause failure. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15 (4): 403–408.
18. Rauzy A. A Benchmark of Boolean Formulae. From http://iml.univmrs.fr/~arauzy/aralia/ benchmark.htm.
19. Revelle C S. A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research 2008; 184 (3): 817-848, https://doi.org/10.1016/j.ejor.2006.12.044.
20. Van der Borst M, Schoonakker H. An overview of PSA importance measures. Reliability Engineering and System Safety 2001; 72 (3): 241-245, https://doi.org/10.1016/S0951-8320(01)00007-2.
21. van Heuven van Staereling I, de Keijzer B, Schafer G. The Ground-Set-Cost Budgeted Maximum Coverage Problem. 41st International Symposium on Mathematical Foundations of Computer Science (MFCS 2016). Dagstuhl Research Online Publication Server, 2016
22. Vaurio J. Ideas and developments in importance measures and fault-tree techiques for reliability and risk analysis. Reliability Engineering and System Safety 2010; 95: 99-107, https://doi.org/10.1016/j.ress.2009.08.006.
23. Vesely W, Davis T, Denning R, Saltos N. Measures of risk importance and their applications. Columbus: Battelle Columbus Labs, OH (USA), 1983, https://doi.org/10.2172/5786790.
24. Wu S. Joint importance of multistate system. Computers & Industrial Engineering 2005; 49: 63-67, https://doi.org/10.1016/j.cie.2005.02.001.
25. Wu S, Coolen F. A cost-based importance measure for system components: An extension of the Birnbaum importance. European Journal of Operational Research 2013; 189-195, https://doi.org/10.1016/j.ejor.2012.09.034.
26. Zafiropoulos E P, Dialynas N E. Methodology for the optimal component selection of electronic devices under reliability and cost constraints. Quality and Reliability Engineering International 2007; 23 (8): 885-897, https://doi.org/10.1002/qre.850.
27. Zaitseva E, Levashenko V, Kostolny J. Application of logical differential calculus and binary decision diagram in importance analysis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2015; 17 (3): 379–388, https://doi.org/10.17531/ein.2015.3.8.
28. Zio E. Risk importance measures. In H. Pham (Ed.), Safety and risk modeling and its applications. London: Springer 2011; 151-196, https://doi.org/10.1007/978-0-85729-470-8_6.
29. Zio E, Podofilini L. Accouniting for components interactions in the differential importance measure. Reliability Engineering and System Safety 2006; 91: 1163-1174, https://doi.org/10.1016/j.ress.2005.11.044.
30. Zio E, Podofilini L. The use of importance measures for the optimization of multi-state systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2006; 2: 33-36.

Petar PAVLOVIĆ
Higher medical and business-technological school of applied studies in Šabac
Hajduk Veljkova 10, 15000 Šabac, Serbia

Dragana MAKAJIĆ-NIKOLIĆ
University of Belgrade
Faculty of Organizational Sciences
Jove Ilića 154, 11000 Belgrade, Serbia

Mirko VUJOŠEVIĆ
University of Belgrade
Faculty of Organizational Sciences
Jove Ilića 154, 11000 Belgrade, Serbia

E-mails: petar.pavlovic@vtssa.edu.rs, makajic-nikolic.dragana@fon.bg.ac.rs, vujosevic.mirko@fon.bg.ac.rs

Article citation info:
GILL A. Optimisation of the technical object maintenance system taking account of risk analysis results. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 420–431, http://dx.doi.org/10.17531/ein.2017.3.13.

Adrian GILL

OPTIMISATION OF THE TECHNICAL OBJECT MAINTENANCE SYSTEM TAKING ACCOUNT OF RISK ANALYSIS RESULTS

OPTYMALIZACJA SYSTEMU OBSŁUGI OBIEKTÓW TECHNICZNYCH Z UWZGLĘDNIENIEM WYNIKÓW ANALIZY RYZYKA*

The article presents the author's original method of optimisation of the technical object maintenance system taking account of risk analysis results. An original form of the objective function was formulated, in which a risk measure model based on two criteria was used. RBM methods were discussed and technical object maintenance methods/strategies were reviewed, with their most important characteristics pointed out. The process of making maintenance-related decisions is armed with procedures based on a risk valuation pattern. The author's original risk valuation pattern was presented and special cases resulting from the use of such patterns were discussed. Dynamic programming was used to solve the problem of optimisation. The author's original mathematical model of the method of optimisation was developed and presented, and its four-stage calculation algorithm was presented in detail. Based on the collected statistical data on damage, hazard analysis and risk assessment procedures were carried out. Using computer implementation of the optimisation model, an experiment in planning the maintenance of the technical objects examined was carried out and the results of the optimisation experiment were presented.

***Keywords**: optimisation, maintenance system, risk, risk based maintenance.*

W artykule przedstawiono autorską metodę optymalizacji systemu obsługi obiektów technicznych z uwzględnieniem wyników analizy ryzyka. Sformułowano oryginalną postać funkcji celu, w której użyto modelu miary ryzyka opartego na dwóch kryteriach. Omówiono metody RBM oraz dokonano przeglądu metod/strategii obsługiwania obiektów technicznych wskazując ich najważniejsze cechy. Podejmowanie decyzji obsługowych uzbrojone jest w procedury oparte na schemacie wartościowania ryzyka. Przedstawiono autorski schemat wartościowania ryzyka i omówiono szczególne przypadki wynikające z użycia takich schematów. Do rozwiązania problemu optymalizacji użyto programowania dynamicznego. Opracowano i przedstawiono autorski matematyczny model metody optymalizacji oraz szczegółowo zaprezentowano jego czteroetapowy algorytm obliczeniowy. Na podstawie zebranych danych statystycznych dotyczących uszkodzeń, przeprowadzono procedury w zakresie analizy zagrożeń i oceny ich ryzyka. Wykorzystując implementację komputerową modelu optymalizacyjnego przeprowadzono eksperyment w zakresie planowania obsług rozpatrywanych obiektów technicznych oraz przedstawiono wyniki eksperymentu optymalizacyjnego.

***Słowa kluczowe**: optymalizacja, system obsługiwania, ryzyko, risk based maintenance.*

1. Introduction

1.1. Preliminary remarks on the subject matter

In numerous organisations, the success of economic projects is conditioned by the level of readiness of technical object elements and their operation systems to perform the tasks imposed on them. The level of readiness of the objects and their components depends, among others, on the effectiveness of the adopted maintenance strategies and procedures. As the author of [30] states: "With the fast development of industry and the highly competitive international market, especially the areas of electronic products, nuclear power, automobile, shipbuilding, and aircraft, cost-effective and accurate maintenance shows increasing importance in improving plant production availability, reducing downtime cost, and enhancing operating reliability". For example – in manufacturing companies, maintenance costs amount to 15-70% of the total cost of production, where most of the maintenance activities function as corrective measures and are implemented primarily in situations particularly requiring their application [34].

In order to improve maintenance activities and decrease maintenance costs, maintenance process models have been under study in a broad scope since 1950 [39]. Thanks to this, numerous concepts or strategies of implementing these processes were created. In principle, they can be divided into two groups: corrective maintenance (CM) and preventive maintenance (PM). Corrective maintenance is performed after damage is identified and its aim is to bring the element and the object to a condition in which they will be able to perform the required functions. This introduces certain limitations, however. For example, according to [30]: "corrective maintenance at its best should be utilized only in non-critical areas where capital costs are small, consequences of failure are slight, no safety risks are immediate, and quick failure identification and rapid failure repair are possible".

Preventive maintenance is performed according to a predetermined schedule or the number of work units (NWU) without previous examination of the condition of the object's elements. If the effects of a poor condition of the object are significantly greater than the effects/costs of preventive activities and it is possible to observe the course of ageing of the object's elements, it is reasonable to use a PM-based strategy [27]. Moreover, as indicated by the author of [33], maintaining a high level of device reliability is achieved, among others, by maintenance work consisting in anticipatory replacement of elements at risk of damage. The advantage of applying PM is that it can be performed at scheduled dates, for example during breaks between the tasks performed. This lowers the risk of interruptions during the per-

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

formance of tasks [27]. The authors of [30, 31] claim in general that a significant advantage of some of the PM-type methods is the creation of conditions for controlling object operation processes.

In the group of PM-type maintenance strategies, condition based maintenance (CBM) and predetermined maintenance can be distinguished. CBM is performed after previous verification or monitoring of the parameters of the object's operation. Monitoring may be continuous or carried out according to a specific schedule. In practice, however, effective application of CBM creates certain challenges. Firstly, its initiation is expensive. Hardware costs are often relatively high. In such case, it is important to decide whether the given element is significant enough to justify the investment. Secondly, making optimal maintenance decisions based on CBM is not always easy to achieve due to variables such as the complexity of the environment, the internal structure of the object, damage mechanisms not entirely known, etc. [30].

According to other authors (e.g. [36]), the following maintenance strategy categories can also be distinguished: *age replacement policy, block replacement policy*, *periodic preventive maintenance policy*, *failure limit policy*, *sequential preventive maintenance policy*, *repair cost limit policy*, *repair time limit policy*, *repair number counting policy*, *reference time policy*, *mixed age policy*, *group maintenance policy*, *opportunistic maintenance policy*, etc.

Moreover, the strategies can be divided, depending on the homogeneity of a set of object elements, into maintenance strategies for homogeneous and heterogeneous elements.

Regardless of the classification, it was observed that the following criteria are usually adopted when developing maintenance strategies [13, 40]:

- maximisation (in the long term) of the technical objects' readiness to perform the tasks imposed on them;
- minimisation of the operating costs of technical objects in the long term;
- maximisation of the technical objects' readiness to perform the tasks imposed on them and minimisation of the average operating costs of these objects in the long term.

The characteristics of technical object maintenance strategies indicated above were used to develop the method presented in this article.

1.2. Risk Based Maintenance

In a modern approach to technical object management, an approach based on the so-called risk management principles is used. As stated by Khan [25], the development of such methods occurred within the last dozen or so years. They are called risk based maintenance (RBM) methods. In various studies from the years 2003-2010 (e.g. [8, 12, 10, 11, 25, 28, 37]), examples of the application of RBM to various types of objects and their elements can be found. Currently (2011-2016), examples of such studies include: [2, 6, 15, 16, 17, 29, 38].

RBM finds application in particular in transport system objects [21]. Their damage or operating errors may generate hazards whose activation is manifested in serious losses – injuries, deterioration of health or loss of human life, considerable material losses, damage to the natural environment or loss of reputation. We can talk about particular importance/significance of such objects or their elements, justifying investments. The problem of the significance of object elements in CBM-type strategies was mentioned in the introduction (chapter 1). In such objects, the need to apply RBM was additionally forced by legislation. As e.g. Zio puts it [40]: "Obviously, occupational and public safety, environmental and other requirements must be satisfied as enforced by regulations". For example, in the EU railway transport system, it results from the CSM (common safety method) documents standardising the requirements and methods connected with the safety of the EU railway system. In accordance with these documents, railway companies, infrastructure administrators, and all the entities introducing changes to the railway system are responsible for maintaining the risk of all the identified hazards at levels below the non-acceptable risk category.

Some conditions of object elements should therefore be examined as the so-called hazard sources (also called hazard factors or risk factors). The term hazard source (HS) can be understood as formations, e.g. physical, chemical, biological, psychophysical, organisational or personal, whose presence in the given area of analysis, condition, properties, etc. are a source of hazard [21].

Identification of the HS may occur at consecutive stages of the objects' life cycle. On this basis, the so-called hazards (H) are formulated, which makes it possible to assess the impact of the damage to these objects on human life and health, the natural environment, and technical systems. This impact is expressed in units of risk level attributed to each hazard, adopting the appropriate risk model. A combination of the level of possibility or probability of hazard activation/materialisation and the level of effects expressed in losses or damage caused by the event is usually adopted. There are numerous examples of such an approach, e.g. in [1, 3, 4, 7, 10, 16, 35]. It should be noted, however, that risk models may include many other components. They may, for example, take account of human errors in the maintenance processes, which was demonstrated, among others, by Hammeed in [15].

In this context of the problem, the organisation of technical object maintenance becomes particularly important. It can and should be treated as a means to achieve the acceptable or at least tolerable level of risk. The most effective, and at the same time most recommended ways of reducing hazard risk are those eliminating the HS. Proper organisation of maintenance meets this condition and in this sense, it may be treated as a tool or means of risk reduction.

Risk reduction measures are usually organised in the form of systems, mainly of a technical nature (e.g. alarm devices, physical covers, protection systems), but the use of measures of an organisational nature (e.g. a team of people acting according to predetermined procedures) is equally effective. The legitimacy of the use of organisational risk reduction solutions is pointed out, among others, by the author of [6], who emphasises the relatively low cost of their implementation.

The maintenance organisation method presented in the article may be rated among RBM methods and risk reduction measures. The RBM concept combines two types of issues, i.e. the issue of object maintenance and the issue of risk assessment. The implementation of risk assessment, also called risk evaluation, consists in verifying (by comparison, valuation) the risk category/class (acceptable, tolerable, non-acceptable) to which the risk specified during the hazard risk analysis belongs. The algorithm of risk evaluation is based on the results of calculations done with the use of the adopted model and is performed according to two procedures: pointing out the risk acceptability and risk valuation areas. It should be added that the task of all the undertakings is to achieve the level of risk in the acceptable risk category or – at least – the tolerable risk category, and to reduce it to the level below the non-acceptable risk category.

The fundamental problem is the appropriate combination of the issues of maintenance and risk assessment, and in particular, the development of proper risk measures and their use in streamlining maintenance. The analyses carried out, among others, in [10], demonstrated that risk conditioning in procedures connected with technical object maintenance, in particular railway vehicle maintenance, can be achieved by:

- the decision criterion – formulation of the form of the objective function of the decision-making problem based on the risk model;
- components of the objective function – taking account of the component concerning the risk of damage to technical objects

or the so-called penalty function connected with the effects of damage to these objects in the formulated objective function;
– limiting the acceptable solution area by adopting the appropriate scope of variability of the decision variables.

It was assumed that the aim of this paper is to develop and present an RBM method consisting in the optimisation of the maintenance system of any technical object.

2. The concept and principles of the presented method of optimisation

In order to solve the problem, it is necessary to formulate the objective function of the decision-making problem based on the risk model or taking account of the component concerning risk. The relevance of this approach is also demonstrated by other authors (including [35]).

The form of the objective function was therefore based on the CURR (cost per unit risk reduction) index, whose model corresponds to both the possibilities referred to above. The description of the CURR index for the so-called risk control option (RCO) was presented, among others, in [26], and this is where its mathematical notation can be found, which is as follows:

$$k_{\Delta R^x;L}^{OSR} = \frac{\Delta K_L^{RCO}}{\Delta R_L^{x;RCO}}, \tag{1}$$

where:

ΔK_L^{RCO} – discounted increase in annual costs connected with the implementation of the given RCO relative to the base option within period L (L may be different than just the expected life cycle of the object),

$\Delta R_L^{x;RCO}$ – risk reduction for loss x within period L after the implementation of the RCO.

The use of the PM-type strategy primarily requires the specification of a certain threshold value of the number of work units for each of the object's elements after which the element becomes damaged. It is expressed in working time units, service life units, the number of starts, etc. The value may be determined e.g. based on the threshold value of tolerable risk.

With the *threshold values of the number of work units of the elements* at hand, one may search for the optimal number of units between object maintenances. The number of work units between object maintenances will therefore be a decision variable in the optimisation model. It was marked as l.

In the case of RBM methods, *the number of work units between object maintenances* is additionally conditioned by the expected value of risk reduction. As the authors of [15] put it: "Shutdown interval is one of the most important factors in determining an effective inspection and maintenance policy. In case if the shutdown inspection and maintenance interval is too short, object shutdown time and production loss along with the inspection and maintenance cost will be too high, vice versa if the shutdown interval is too long, the production loss and inspection and maintenance cost will be low but the risk exposure will be high" [15].

Therefore, the problem lies in the determination of the optimal number of work units between object maintenances and thus obtaining the optimal numbers of work units between object element maintenances.

Many problems of streamlining operation systems, including maintenance problems, may be boiled down to solving the tasks of static optimisation. They are usually tasks of non-linear optimisation with limitations. A review of the models of optimisation of technical object maintenance was presented, among others, by the author of [5], and examples of such models are presented, among others, by the authors of [35].

So treating the maintenance system as a point in a certain multidimensional Euclidean space, one may divide the set of this point's coordinates into two subsets: the variables called parameters, which are taken as constants in the process of configuring the maintenance cycle, and the decision variables whose values are modified. Taking into consideration the form of the selected CURR index, the threshold values of the number of work units and the costs of activities concerning the so-called risk handling procedures (taking an active attitude towards the identified hazards) were adopted as the parameters of the optimisation model. In the presented problem, it was assumed that such attitude involves preventive maintenance activities consisting in restoring the usability of the appropriate elements of the technical object under maintenance.

The number of work units between object maintenances and the number of work units between element maintenances were adopted as the decision variables. It is assumed that the number of work units between element maintenances is a multiple of the number of work units between object maintenances. The number of work units between element maintenances was marked as l_i.

Another issue concerns the link between the components of the risk model which was included in the objective function and the decision variables, i.e. the expression of the components of the risk model with the use of the decision variables. For this purpose, the generalised risk model presented by Kadziński, among others in [21], will be used. Subsequent versions of this model are published, among others, in [23].

The risk model for each hazard is a function of the components which are the result of separate decisions made based on the analysis according to m criteria K_k ($k = 1,2,...,m$). In accordance with the definition of risk, each of the m analysis criteria has to be such that component $r_k(z_i)$ ($k = 1,2,...,m$), which is the result of a decision made according to this criterion, should belong to the group of components expressing the possibility of the activation of hazard z_i ($i = 1,2,...,n$) or the size of the potential damage/losses resulting from its activation [21]. Thus defined, the risk may be the basis for formulating the criteria of optimisation of both the periods and scopes of preventive refurbishment [3, 35].

When the levels of all the risk components are determined, the total risk R of hazard z_i ($i = 1,2,...,n$) may be notated as follows:

$$R(z_i) = f_1\left(r_1(z_i), r_2(z_i), \ldots, r_m(z_i)\right), \quad i = 1,2,\ldots,n \tag{2}$$

where:

n – the number of identified hazards,
m – the number of risk analysis criteria,
z_i – i-th hazard from the set of identified hazards,
$r_k(z_i)$ – k-th component of risk z_i within the scope of the k-th risk analysis criterion.

And so the value of RT (Risk Treatment) – the index of the procedure of handling the risk of hazard connected with the condition of the object's elements shall be the function of two components:

$$RT = f(c_i, \Delta R_i) \tag{3}$$

where:

c_i – the cost of the procedure of handling the risk of the i-th hazard (generated by the condition of the i-th object element),

ΔR_i – the value of the reduction of the risk of the *i*-th hazard achieved as a result of avoiding damage to the *i*-th object element.

Function *f* defined in dependence (3) is adopted as the objective function of the decision-making problem.

As indicated before, the achievement of a level of risk below the non-acceptable risk category is satisfactory. It is best, however, if achieving and maintaining risk at the level of the acceptable risk category is possible. Between the two, there is the area of the tolerable risk category. Such division of the risk space is often adopted e.g. in the widely applied ALARP concept.

So let us assume that the reaction of the optimisation model will occur when the value of the risk of the *i*-th hazard R_i is in the area of the tolerable risk category. This means that ΔR_i, i.e. changes in the value of hazard risk reduction will be included in the following range:

$$0 \le \Delta R_i \le R_i^{GT} - R_i^{DT} \tag{4}$$

where:

R_i^{DT} – lower limit of the area of the category of tolerable risk of the *i*-th object element,

R_i^{GT} – upper limit of the area of the category of tolerable risk of the *i*-th object element.

In accordance with the principles concerning the risk model and dependence (2), among the components $r_k(z_i)$ ($k = 1,2,...,m$), at least one belongs to the group of components expressing the possibility of hazard activation. In the presented optimisation problem, the hazard connected with the condition of the object elements resulting from the number of work units (time) is subject to risk assessment. Thus, at least one (*j*-th) of the components of the risk of the *i*-th hazard is dependent on the working time of the object element with which the hazard is connected.

One of the possible ways of expressing the possibility of hazard activation is the probability of the event where the condition of the working element will require the performance of maintenance activities after working for the *l* number of work units. This can be notated as follows:

$$r_{ij} = F_i(l) = P(L_i < l) \tag{5}$$

where:

L_i – random variable expressing the number of work units until the damage of the *i*-th object element,

r_{ij} – *j*-th component of the hazard risk of the *i*-th object element dependent on the *l* number of work units.

For the purpose of this model, it was assumed that a single hazard is connected with damage to the *i*-th object element. Moreover – taking account of dependence (5) – is it suggested that the risk of this hazard should be notated as follows:

$$R_i = f_2\left[r_{i1}, r_{i2}, \ldots, r_{im-1}, F_i(l)\right], \quad i = 1,2,\ldots,N \tag{6}$$

where:

N – the number of identified hazards equal to the number of object elements (table 4),

R_i – the risk of hazard connected with the condition of the *i*-th object element requiring maintenance after the element having worked for *l* work units.

For value R_i of function f_2 equal to R_i^{GT}, there is a possibility of determining l_i^{GT} – a threshold value of the number of work units of the *i*-th object element (example – table 4). The manner of determining this number was presented in [12]. If the threshold value R_i^{GT} is unknown, the number l_i^{GT} may be determined e.g. with the use of the model presented in [14].

It is also possible to determine a certain number l_i^{DT}, i.e. the threshold value of the number of work units of the *i*-th object element, after the exceeding of which, it is reasonable to perform maintenance. It is not justified, however, to plan this maintenance before l_i^{DT}, as the risk connected with the condition of the object element is acceptable at the time. Therefore, the range $\left\langle l_i^{DT}; l_i^{GT} \right\rangle$ determines a practical scope of the value of the NWU of the *i*-th element of the object in which the decision about the need to perform maintenance is made, i.e.:

$$l_i^{DT} \le l_i \le l_i^{GT} \tag{7}$$

where:

l_i – the number of work units between maintenances of the *i*-th object element.

Value l_i^{DT} will serve to determine ΔR_i – reduction of the risk of the *i*-th hazard, achieved as a result of having anticipated damage to this element. Using dependences (4), (6), and (8), value ΔR_i shall be determined according to the following function:

$$\Delta R_i = f_3(R_i) = \begin{cases} f_2\left[F_i(l) \cong 0\right] - R_i^{\min} & dla \quad l_i \le l_i^{DT} \\ f_2(l = l_i) - R_i^{\min} & dla \quad l_i^{DT} < l_i \le l_i^{GT} \end{cases}, \tag{8}$$

where $R_i^{\min}$ is the minimum value of risk and the beginning of the scope of the acceptable risk category. The value of risk is practically always larger than zero.

And so taking account of formulas (2), (6), and (8), in which function f_1, f_2, f_3 was defined, the objective function *f* may be notated as follows:

$$RT = f_4[c_i, f_3(R_i)] \Rightarrow RT = f\ \{c_i, f_3[f_2(r_{i1}, r_{i2}, \ldots, r_{im-1}, F_i(l))]\} \tag{9}$$

Let us assume further that a single maintenance of the object consists in complete renewal of the operation potential of one or several of its elements at the same time. Such maintenance restores all the parameters to such a condition that the refurbished element may be treated as new. So there is a certain value l^* – the value of the optimal number of work units between object maintenances which minimises the value of the objective function (9), i.e.:

$$RT = f\ (l^*) = \min \tag{10}$$

The manner of determining this value is presented in the detailed model described further on in the article.

3. The mathematical model

At an initial stage of calculations, the $i\ (i = 1, 2, \ldots N)$ index assigned to the individual elements of the object shall be understood as the element's identifier resulting from a position series created according to the increasing l_i^{GT}, i.e.

$$l_1^{GT} \le l_2^{GT} \le \ldots \le l_N^{GT}\ . \tag{11}$$

In accordance with dependence (11), the first maintenance of the element with identifier $i-1$ will be performed earlier than that of element *i*, which means that $l_1 \le l_2 \le \ldots \le l_N$.

In the PM-type strategies, maintenances may be performed at fixed intervals. Moreover, the maintenance activities of the so-called lower order (usually occurring earlier and more frequently) are included within the scope of the maintenances of the higher order. This leads to the occurrence of certain dependencies between the values of the numbers of work units between maintenances of object elements. The values of these numbers, occurring later in the maintenance cycle, are a multiple of the numbers of work units between maintenances occurring earlier. This was notated with the introduction of the multiplicity factor a_i assuming values from the set of positive integers:

$$a_i = \mathrm{int}\left(\frac{l_i}{l_{i-1}}\right),\quad a_i \in \mathbf{C}\ \ (i = 2,\ldots,N)\ . \tag{12}$$

Value a_i was used to express the number of work units between maintenances of elements with identifiers $i = 2, 3,\ldots,N$:

$$\begin{aligned} l_2 &= a_2 l_1; \\ l_3 &= a_3 l_2 = a_3 a_2 l_1; \\ &\ldots \\ l_i &= a_i a_{i-1} \ldots a_3 a_2 l_1; \\ &\ldots \\ l_N &= a_N a_{N-1} \ldots a_{i-1} \ldots a_3 a_2 l_1. \end{aligned}\ , \tag{13}$$

where:

i – object element identifier resulting from the sequence in the position series created from value l_i^{GT} (table 4),

$a_i\ (i = 1, 2,\ldots,N)$ –multiplicity factor between the numbers of work units between maintenances of object elements ($a_i \in \mathbf{C}$).

And so a certain vector $\mathrm{L} = [l_1, l_2, \ldots, l_N]$ of work units between maintenances of object elements which minimises the value of the objective function (9) is searched for. It will be called the vector of the decision variables of the optimisation model. Using dependences (10) and (13), one may notate this vector as follows:

$$\mathrm{L} = [l_1, a_2, a_3, \ldots, a_N]\ . \tag{14}$$

And taking into consideration formulas (1) and (3) and dependences (13) and (14), the objective function *f* may be initially expressed as follows:

$$RT = f(\mathrm{L}) = \sum_{i=1}^{N} \frac{c_i}{\Delta R_i}\ , \tag{15}$$

with two limitations resulting from dependences (7) and (13):

$$0 < l_1 \le l_1^{GT}\ , \tag{16}$$

$$0 < l_1 \cdot \prod_{v=2}^{i} a_v \le l_i^{GT}\ ,\ \ i = 2,3,\ldots,N\ . \tag{17}$$

Limitations (16) and (17) are linear functions, while function (15) is non-linear. The problem of determining the optimal number of work units between object maintenances may be solved using the methods of non-linear mathematical or dynamic programming [9], which is presented below.

If $q_i(l_i)$ is used to denote a component of the *RT* index determined for the *i*-th hazard (of the *i*-th object element), i.e.:

$$q_i(l_i) = \frac{c_i}{\Delta R_i}\ ,\ \ i = 1,2,\ldots,N\ , \tag{18}$$

the objective function *f* may be expressed generally, in shortened form, as follows:

$$f(l_1, a_2, a_3, \ldots, a_N) = \sum_{i=1}^{N} q_i(l_i)\ . \tag{19}$$

Using RT^* to denote the minimum value of the index of the procedure of handling the risk of hazards concerning *N* object elements and taking account of dependence (10), dependence (19), and the form of vector L, one may note that:

$$RT^* = f\left(l_1^*, a_2, a_3, \ldots, a_N\right) = \min_{l_1 \in \Lambda_1, \ldots, l_N \in \Lambda_N} \left\{ \sum_{i=1}^{N} q_i(l_i) \right\}. \tag{20}$$

where Λ_i are sets of the numbers of work units between object element maintenances, containing discrete values of variables l_i $(i = 1,2,3,\ldots,N)$. The problem of discretisation of the scope of values l_i was presented in chapter 4 of the article.

Component $q_1(l_1)$ constitutes an invariable component of each of the values of acceptable solutions, and so it can be excluded from below the symbol *min*:

$$f(l_1, a_2, a_3, \ldots, a_N) = q_1(l_1) + \min_{l_2 \in \Lambda_1, \ldots, l_N \in \Lambda_N} \left\{ \sum_{i=2}^{N} q_i(l_i) \right\}. \tag{21}$$

Moreover, if it is assumed that the values of the components of the *RT* index connected with the hazards concerning the condition of the elements with identifiers $i = 2,3,\ldots,N$ are equal:

$$\min_{l_2 \in \Lambda_1, \ldots, l_N \in \Lambda_N} \left\{ \sum_{i=2}^{N} q_i(l_i) \right\} = u_2(l_2)\ , \tag{22}$$

then taking account of formula (21), one may determine a certain function $w_1^*(l_1)$ – of the minimum RT values connected with the hazards concerning the condition of all the object elements $(i = 2, 3,\ldots,N)$ determined for any value $l_1 \in \Lambda_1$:

$$w_1^*(l_1) = q_1(l_1) + \min_{l_2 \in \Lambda_2} u_2(l_2) \,. \tag{23}$$

The optimal value of the *RT* index for *N* object elements will therefore equal::

$$f^* = \min_{l_1 \in \Lambda_1} w_1^*(l_1) \,. \tag{24}$$

In accordance with the methodology of dynamic or mathematical programming presented e.g. in [32], determining the minimum *RT* value for *N* technical object elements should begin with determining the minimum value of the *RT* value for the element with identifier $i = N$:

$$u_N(l_N) = \min_{l_N \in \Lambda_N} q_N(l_N) = \min_{l_N \in \Lambda_N} \frac{c_N}{\Delta R_N} \tag{25}$$

and then for subsequent elements with identifiers $i = N-1, N-2, ..., 1$.

The *RT* value connected with the element with identifier $i = N-1$ depends on the value of the *RT* component determined for the element with identifier $i = N$, i.e.:

$$\begin{gathered} u_{(N-1)}(l_{(N-1)}) = \min_{l_{N-1} \in \Lambda_{N-1}} \left[q_{(N-1)}(l_{(N-1)}) + u_N(l_N) \right], \\ \ldots \\ u_i(l_i) = \min_{l_i \in \Lambda_i} \left[q_i(l_i) + u_{i+1}(l_{i+1}) \right], \\ \ldots \\ u_3(l_3) = \min_{l_3 \in \Lambda_3} [q_3(l_3) + u_4(l_4)] \,, \\ u_2(l_2) = \min_{l_2 \in \Lambda_2} [q_2(l_2) + u_3(l_3)] \,. \end{gathered} \tag{26}$$

The number of work units between maintenances of element l_1, for which the *RT* value turned out to be the smallest, is also the optimal number of work time units between object maintenances. Based on this number and using dependence (13), one may determine the remaining optimal values $l_i^* \, (i = 2, 3, ...N)$.

4. The scope of calculations and details of the mathematical model

Stage 1 – Adoption of the risk model and risk valuation model

For the purpose of the presentation of the scope of necessary calculations, using the results of works [21, 24], a detailed model of a risk analysis process for hazards connected with the condition of object elements was adopted. Within the model, two ($k = 1, 2$) analysis criteria were adopted:

- K_1 – criterion of damage/losses suffered as a result of hazard activation,
- K_2 – criterion of the possibility of hazard activation.

An example of a pattern of quantification of the levels of damage/losses resulting from hazard activation, within the scope of the first risk analysis criterion, was presented in table 1.

Table 1. An example of a pattern of quantification of the levels of damage/losses resulting from hazard activation

j	Level of damage/losses	Characteristics of damage/losses suffered as a result of hazard activation
1	green	Minor injuries of object users and/or co-users of the space / low level of material damage
2	blue	Injuries of object users and/or co-users of the space / measurable level of material damage
3	yellow	Serious injuries of object users and/or co-users of the space / significant level of material damage
4	orange	Single fatalities among object users and/or co-users of the space / high level of material damage
5	red	Numerous fatalities among object users and/or co-users of the space / very high level of material damage

Source: prepared based on [23]

The values of the risk component within the scope of the second analysis criterion are determined in accordance with dependence (5). The basis for calculations is the data on object operation, including above all data on damage to elements. Knowledge of the wear processes which may lead to damage of these elements is also essential [35]. Such data for a sample object is presented in table 3.

In the detailed model of maintenance system optimisation, the following set of risk analysis criteria significance measures was adopted:

$$A = \{2,1\}, \tag{27}$$

and the elements of the set of measures of the second risk component were assigned the following set of values:

$$\Omega_k = \left\{ \varpi_{ik}^{(j)} \right\} = \{0,250; 0,375; 0,500; 0,625; 1,000\}, i = 1,2,...,n; j = 1,2,...,5; k = 1. \tag{28}$$

The last issue is the selection of the form of function f_2 – dependence (6), which makes it possible to determine value R_i. It may be determined in the form of a mathematical function, tabularly, with a graph, verbally or otherwise. Apart from the levels of risk components, risk analyses also take into account the significance measures of the analysis criteria, and one of the most frequent dependences is (29):

$$R_i = \prod_{k=1}^{m} \alpha_{ik} \cdot r_{ik}, \; i = 1, 2, \ldots, N \,, \tag{29}$$

or dependence (30) taking into account the results of hazard risk analysis according to two criteria and the significance measures of the hazard risk analysis criteria:

$$R_i = \sum_{k=1}^{2} \alpha_{ik} \cdot r_{ik}, \; i = 1, 2, \ldots, N \tag{30}$$

where α_{ik} are the significance measures of risk components within the scope of the *k*-th criterion of risk analysis. Such form of the risk function is applied in many known methods. For example, in the risk score method risk model and the failure mode and effects analysis method, three components ($k = 3$) and a set of measures of their significance $A = \{1; 1; 1\}$ are applied.

So using dependences (9), (15), and (19), the objective function *RT* values shall be expressed in formula (31):

$$RT = \sum_{i=1}^{N} \frac{c_i}{\alpha_m B_i \cdot \prod_{k=1}^{m-1} \alpha_{ik} \cdot r_{ik} - R_i^{\min}}, \tag{31}$$

where B_i is a component of the objective function dependent on the distribution of probability describing the number of work units until the damage of the *i*-th element. These may be typical probability distributions. It might e.g. be initially assumed that element maintenance occurs in the period of their age-related damage, which would be very advantageous from the point of view of the effectiveness of the maintenance system operation. For in this period, damage is usually characterised by normal probability distribution. In such case, the form of the objective function *f* may e.g. be as follows:

$$RT = \sum_{i=1}^{N} \frac{c_i}{\frac{\alpha_m}{\sigma_i \cdot \sqrt{2 \cdot \Pi}} \cdot \int_{-\infty}^{l_i} e^{-\frac{(l_i - \mu_i)^2}{2 \cdot \sigma_i^2}} \cdot dl \cdot \prod_{k=1}^{m-1} \alpha_{ik} \cdot r_{ik} - R_i^{\min}}, \tag{32}$$

or otherwise:

$$RT = \sum_{i=1}^{N} \frac{c_i}{\alpha_m \cdot F_i^{N(\mu;\sigma)}(l_i) \cdot \prod_{k=1}^{m-1} \alpha_{ik} \cdot r_{ik} - R_i^{\min}} \tag{33}$$

where:

$F_i^{N(\mu;\sigma)}(l_i)$ – value of the cumulative distribution function of the number of work units until the damage of the *i*-th object element with normal distribution, determined for the number of work units between element maintenances l_i,

μ_i – expected value of the number of work units until the damage of the *i*-th object element,

σ_i – standard deviation of the number of work units until the damage of the *i*-th object element.

Stage 2 – Identification of hazards generated by damage to object elements

Hazard identification is a separate issue and due to its complexity, it will not be raised in this article. A proposal of the implementation of a hazard identification process can be examined in the following publications of the author of this article: [18-20]. The result of the hazard identification process is the so-called hazard record.

Stage 3 – Estimation and valuation of the risk of hazards generated by the condition of object elements

In the presented optimisation model, the risk level space was divided into three areas. The pattern of this valuation is presented in table 2.

When the risk level is included in the area of the non-acceptable or tolerable risk category, risk handling procedures, i.e. maintenance activities should be implemented.

Due to the adopted ranges of risk levels (table 2), two extreme cases concerning the determination of l_i^{GT} based on value R_i may occur in the calculation process. The first case consists in the fact that the value of risk R_i will always be in the A area, i.e. the acceptable risk category, regardless of value r_{i2} – risk component expressing the possibility of hazard activation. This happens when the level of damage/losses resulting from hazard activation is very low (e.g. level 1 – "green").

Table 2. Pattern of risk valuation in the model of optimisation of the technical object maintenance system

Range of values of risk measures	Range of risk levels	Name of the risk category area and its symbol
$\left[R_i^{\min}, R_i^{DT}\right)$	[0.50; 1.40)	Acceptable risk category area – symbol A
$\left[R_i^{DT}, R_i^{GT}\right]$	[1.40; 1.60]	Tolerable risk category area – symbol T
$\left(R_i^{GT}, R_i^{\max}\right]$	(1.60; 3.00]	Non-acceptable risk category area – symbol NA

Source: own elaboration based on [21]

The second case consists in the fact that the value of risk R_i will always be in the NA area, i.e. the non-acceptable risk category, also regardless of value r_{i2}. This is the case when the first of the components of risk r_{i1} assumes the highest values of damage/loss measures, e.g. values at the "red" level (table 2).

The threshold values of the number of work units of elements may be obtained based on values R_i^{DT} and R_i^{GT}. And so r_{i2}^{DT} shall be a certain threshold value of the second component obtained based on value R_i^{DT}. Assuming that risk R_i is given with dependence (30), it may be notated that:

$$r_{i2}^{DT} = \frac{R_i^{DT} - \alpha_{i1} \cdot r_{i1}}{\alpha_{i2}} \tag{34}$$

and using dependence (5):

$$r_{i2}^{DT} = F_i^{N(\mu,\sigma)}(l_i^{DT}), \tag{35}$$

hence:

$$F_i^{N(\mu,\sigma)}(l_i^{DT}) = \frac{R_i^{DT} - \alpha_{i1} \cdot r_{i1}}{\alpha_{i2}} \Rightarrow l_i^{DT} = F_i^{-1N(\mu,\sigma)}\left(\frac{R_i^{DT} - \alpha_{i1} \cdot r_{i1}}{\alpha_{i2}}\right). \tag{36}$$

Knowing that the cumulative distribution function of the number of work units until the damage of the *i*-th object element has normal distribution, it may be assumed that:

$$F_{i(\max)}(l) = F_i^{N(0;1)}(3) \tag{37}$$

In the first case, concerning the determination of l_i^{GT}, when the value of risk R_i will always be in area A, there is no need to "hurry" with performing the maintenance of the object element. In the second case, when the value of risk R_i will always be in area NA, element maintenance should be performed as soon as possible. Within the scope of the detailed model and the adopted form of function (15) or (20), and in particular in the form of the cumulative distribution function, a solution in the following form is suggested:

$$l_i^{DT} = \begin{cases} \mu_i - 3\sigma_i & dla \quad r_{i2} \cong 0 \\ F_i^{-1}\left(\dfrac{R_i^{DT} - \alpha_{i1} \cdot r_{i1}}{\alpha_{i2}}\right) & dla \quad 0 < r_{i2} < 1 \\ \mu_i + 3\sigma_i & dla \quad r_{i2} \cong 1 \end{cases}$$

$$l_i^{GT} = \begin{cases} \mu_i - 3\sigma_i & dla \quad r_{i2} \cong 0 \\ F_i^{-1}\left(\dfrac{R_i^{GT} - \alpha_{i1} \cdot r_{i1}}{\alpha_{i2}}\right) & dla \quad 0 < r_{i2} < 1 \\ \mu_i + 3\sigma_i & dla \quad r_{i2} \cong 1 \end{cases} \tag{38}$$

The cases referred to above do not have to occur with differently adopted risk valuation patterns.

Stage 4 – Determining the ranges of values of the decision variables of the optimisation model

Determining the ranges of values of the decision variables of the model primarily serves the execution of calculations in the form of a simulation. This is required by the research problem posed, and in particular the non-linear form of the objective function.

The value of variable l_1 is a real number from the range of $\left\langle l_1^{DT}; l_1^{GT} \right\rangle$. And so discretisation of the range of values l_1 is introduced, i.e. a certain step of simulations/calculations Δl is adopted and it is assumed that $l_i^{DT} = \Delta l$. Δl can also be interpreted as the accuracy of the calculations.

Discretisation of the range of values l_1 and the multiplicity between object maintenances (dependence (13)) lead to the formation of distinctive sets of decision variables, i.e. sets of possible numbers of work units between element maintenances.

And so Λ_i shall be the sets of possible numbers of work units between element maintenances which are defined as follows:

- for $i = 1$

$$\Lambda_1 = \left\{ l_{1(j)} \right\}, \; j = 1, 2, ..., m_1, \text{ where } l_{1(j)} - l_{1(j-1)} = \Delta l \tag{39}$$

where:

$l_{i(j)}$ – is the j-th possible value of the number of work units between maintenances of the i-th element.

j – is the identifier of the next possible value of the number of work units between maintenances of the i-th element, which belongs to set Λ_1,

m_1 – is the size of set Λ_1 equal to:

$$m_i = \text{int}\left(\frac{l_1^{GT}}{\Delta l}\right), \tag{40}$$

- for i > 1

$$\Lambda_i = \left\{ l_{i(k)} \right\}, \; k = 1, 2, ..., n_i \tag{41}$$

where:

n_i – is the size of set Λ_i dependent on value $l_{1(j)}$ as follows:

$$n_i = \text{int}\left(\frac{l_i^{GT}}{l_{1(j)}}\right), \tag{42}$$

and:

$l_{i(k)}$ – is the k-th possible value of the number of work units between maintenances of the i-th element, such that:

$$\underset{i>1}{\forall}\underset{j}{\forall}\underset{k}{\forall} l_{i(k)} = k \cdot \prod_{v=2}^{i} a_{v-1} \cdot l_{1(j)}, \tag{43}$$

Figure 1 presents, among others, the manner of determining sets Λ_i for four elements of a sample technical object and one of the values $l_{1(j)}$.

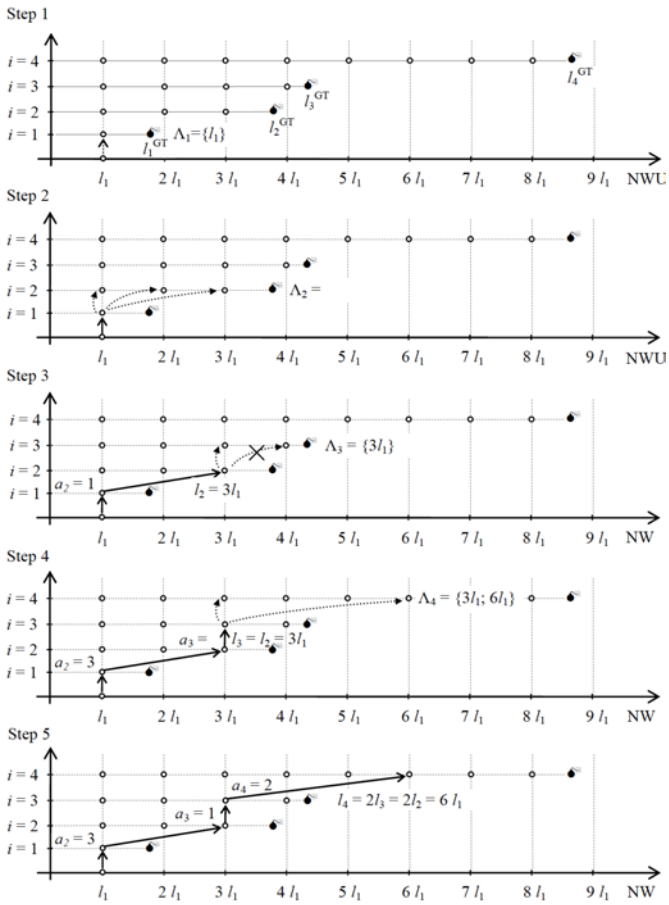


Fig. 1. Diagram of obtaining a sample solution acceptable in the optimisation of the number of work units (NWU) between technical object maintenances for four elements of this object and one of the values l_1

So assuming that $i = 4$, $j = 1$ and that the decision variables assume values as presented in figure 1, i.e.: $a_1 = 1$, $a_2 = 3$, $a_3 = 1$, $a_4 = 2$, sets Λ_i will be as follows:

$$\Lambda_1 = \left\{ l_{1(1)} \right\}$$

$$\Lambda_2 = \left\{ l_{2(1)}; l_{2(2)}; ...; l_{2(n_2)} \right\} = \left\{ a_1 l_{1(1)}; 2a_1 l_{1(1)}; ...; n_2 a_1 l_{1(1)} \right\} = \left\{ l_{1(1)}; 2l_{1(1)}; 3l_{1(1)} \right\}$$

$$\Lambda_3 = \left\{ l_{3(1)}; l_{3(2)}; ...; l_{3(n_3)} \right\} = \left\{ a_2 a_1 l_{1(1)}; 2a_2 a_1 l_{1(1)}; ...; n_3 a_2 a_1 l_{1(1)} \right\} = \left\{ 3l_{1(1)} \right\}$$

$$\Lambda_4 = \left\{ l_{4(1)}; l_{4(2)}; ...; l_{4(n_4)} \right\} = \left\{ a_3 a_2 a_1 l_{1(1)}; 2a_3 a_2 a_1 l_{1(1)}; ...; n_3 a_3 a_2 a_1 l_{1(1)} \right\} = \left\{ 3l_{1(1)}; 6l_{1(1)} \right\}$$

Using dependence (26), one may properly notate the dependences of individual components of the RT value. The notation of these values in individual points of the graph presented in figure 1 and for one value $l_{1(j)}$ (e.g. $j = 1$) will be as follows:

- for $i = 1$

$$u_1(l_1 = l_{1(1)}) = q_1(l_{1(1)}) + \min_{l_2 \in \Lambda_2 = \{l_{1(1)};\, 2l_{1(1)};\, 3l_{1(1)}\}} u_2(l_2)$$

- for $i = 2$

$$u_2(l_2 = l_{2(1)} = l_{1(1)}) = q_2(l_{1(1)}) + \min_{l_3 \in \Lambda_3 = \{l_{1(1)};\, 2l_{1(1)};\, 3l_{1(1)};\, 4l_{1(1)}\}} u_3(l_3)$$

$$u_2(l_2 = l_{2(2)} = 2l_{1(1)}) = q_2(2l_{1(1)}) + \min_{l_3 \in \Lambda_3 = \{2l_{1(1)};\, 4l_{1(1)}\}} u_3(l_3)$$

$$u_2(l_2 = l_{2(3)} = 3l_{1(1)}) = q_2(3l_{1(1)}) + \min_{l_3 \in \Lambda_3 = \{3l_{1(1)}\}} u_3(l_3)$$

- for $i = 3$

$$u_3(l_3 = l_{3(1)} = l_{1(1)}) = q_3(l_{1(1)}) + \min_{l_4 \in \Lambda_4 = \{l_{1(1)};\, 2l_{1(1)};\, 3l_{1(1)};\, 4l_{1(1)};\, 5l_{1(1)};\, 6l_{1(1)};\, 7l_{1(1)};\, 8l_{1(1)}\}} u_4(l_4)$$

$$u_3(l_3 = l_{3(2)} = 2l_{1(1)}) = q_3(2l_{1(1)}) + \min_{l_4 \in \Lambda_4 = \{2l_{1(1)};\, 4l_{1(1)};\, 6l_{1(1)};\, 8l_{1(1)}\}} u_4(l_4)$$

$$u_3(l_3 = l_{3(3)} = 3l_{1(1)}) = q_3(3l_{1(1)}) + \min_{l_4 \in \Lambda_4 = \{3l_{1(1)};\, 6l_{1(1)}\}} u_4(l_4)$$

$$u_3(l_3 = l_{3(4)} = 4l_{1(1)}) = q_3(4l_{1(1)}) + \min_{l_4 \in \Lambda_4 = \{4l_{1(1)};\, 8l_{1(1)}\}} u_4(l_4)$$

- for $i = 4$

$$u_4(l_4 = l_{4(k)}) = q_4(l_{1(k)}), \quad k = 1, 2, ..., 8$$

As a result, the first value of component $w_1(l_1 = l_{1(1)})$ of the *RT* index obtained according to the formula presented in figure 1 can be obtained in the following manner:

$$w_1(l_1 = l_{1(1)}) = q_1(l_{1(1)}) + q_2(3l_{1(1)}) + q_3(3l_{1(1)}) + q_4(6_{1(1)}).$$

Another step of the calculations is the determination of value $w_1(l_1)$ for the remaining values $l_{1(j)}$, i.e. for $j = 2, 3, ...m_1$.

5. Application of the optimisation model for a sample object

An example of optimisation was presented for certain elements of a tram. The most frequently damaged elements of various tram systems were selected. Their damage usually causes the tram to stop. The condition of the elements changes as a result of wear and/or it is relatively difficult or unprofitable to monitor the technical condition of these elements.

Information about damage to selected elements of the tram was presented in the form of appropriate characteristics (table 3). It was obtained based on the data (collected for the purposes of the Municipal Transport Company (MPK)) on the times and reasons for the so-called emergency returns to depot. The data is from the period of three years during which the trams were operated in normal working conditions. Their working time expressed in days was selected as a variable characterising the life cycle of the elements. Hypotheses on the form of the probability distributions of this time were verified with the use of the Statistica 12 program, assuming the level of significance of 0.05.

The parameters of the optimisation model were presented in table 4. In order to determine them, the identification of hazards (stage 2, chapter 4) connected with damage to selected elements was performed, and the formulated hazards were provided in column 5 of table 4. The hazards were given identification numbers ("H_") corresponding to identifier *i* of the elements.

Next, guided by the principles of the adopted risk model (stage 1, chapter 4), the values of one of the risk components were adopted (column 6, table 4). This made it possible to determine values l_i^{GT} (column 8, table 4) in accordance with dependence (38) with known function $F_i(l)$, given with the proper probability distribution (table 3).

On account of the complexity of the calculations (stage 4, chapter 4) and a specific character of the optimisation problem, it was decided that a dedicated computer application for the optimisation model should be developed. It makes it possible to do calculations in the form of simulation for theoretically any number of object elements with assumed simulation/calculation step Δl. Figure 2a presents the diagram of the structure of maintenance cycles and a chart of the values of the objective function for the object adopted for the presented

Table 3. Characteristics of selected tram elements

Item	Element name	Descriptive statistics of the element's working time				Characteristics of distribution matching			
		Number of implementations	Number of cars	Average value	Standard deviation	Probability distribution type	Degrees of freedom	Chi-squared statistics value	Statistical significance *p*
1	Resistance lamination sheets of the GBT-373 starter	394	43	87.10914	120.3005	Log-normal	13	19.25974	0.11527
2	Current collector contact shoe	164	55	189.6890	182.3782	Exponential	9	10.28231	0.32812
3	Current converter motor brushes	54	12	68.55556	62.30953	Log-normal	3	1.94371	0.58417
4	Brake lever mechanism	608	35	51.93092	75.09104	Log-normal	14	19.48904	0.14709
5	Universal joint	44	29	221.6364	237.6264	Exponential	8	13.54065	0.09455
6	R15 transmitter	60	25	149.9000	175.7922	Log-normal	3	4.61922	0.20190
7	Door controller	49	25	168.7959	194.0156	Exponential	2	2.99906	0.22324
8	Electronic starting transmitter (EST)	64	37	210.4063	201.1366	Exponential	6	6.75927	0.34368
9	Door mechanism cam adjustment	66	33	232.6667	221.4298	Exponential	4	8.71243	0.06870

Table 4. Parameters of the optimisation model of the technical object maintenance system

Item	i*	Element name / (type of maintenance activity)	ID**	Hazard connected with damage to tram element	Value of the measure of risk component expressing damage/losses after hazard activation r_{i1}	Cost ci of risk handling procedures expressed in time losses [min]	Threshold value of the number of work units for the element l_i^{GT} [days]
1	2	3	4	5	6	7	8
1	5	Resistance lamination sheets of the GBT-373 starter/(replacement)	BO	H5 – risk of tram stopping while in use	0.250	360	448.01
2	4	Current collector contact shoe/(replacement)	SO	H4 – risk of traction network damage	0.375	180	359.86
3	2	Current converter motor brushes/(replacement)	SS	H2 – risk of tram stopping while in use	0.250	60	255.48
4	1	Brake lever mechanism/(adjustment)	MD	H1 – risk of extending the braking distance	0.500	240	70.96
5	9	Universal joint/(replacement)	PK	H9 – risk of tram stopping while in use	0.250	420	934.52
6	6	R15 transmitter/(replacement)	PR	H6 – risk of tram stopping while in use	0.250	120	677.28
7	3	Door controller/(replacement)	SD	H3 – risk of passenger being hit and/or knocked over	0.375	300	320.23
8	7	Electronic starting transmitter (EST)/(replacement)	PE	H7 – risk of tram stopping while in use	0.250	960	813.82
9	8	Door mechanism cam/(adjustment)	KM	H8 – risk of doors not closing properly preventing the tram from departing	0.250	120	896.96

** – identifier of object element resulting from the increasing threshold values of the numbers of work units of elements*

*** – identifier of the element in the maintenance cycle structure diagram*

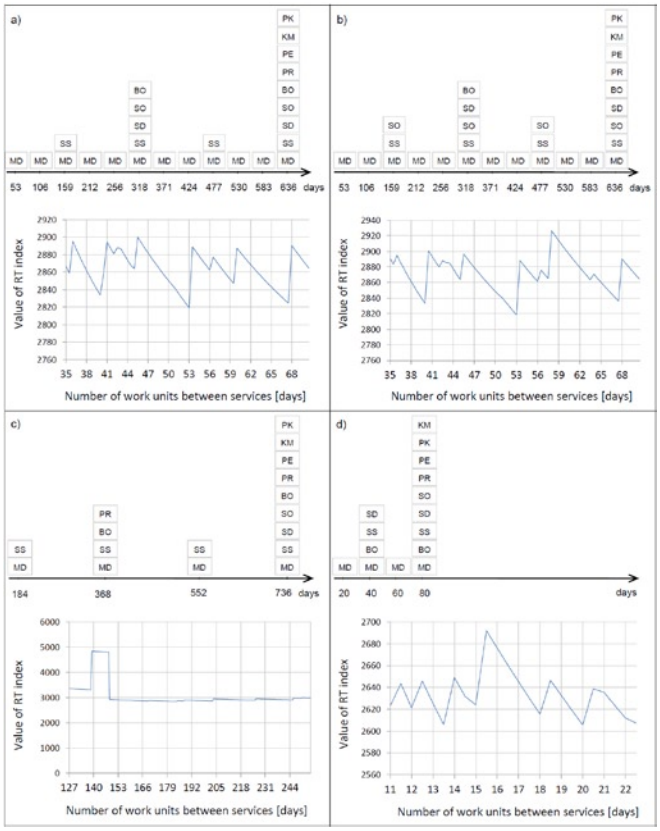


Fig. 2. The results of determining the optimal structure of the maintenance cycle with the use of the optimisation model for different variants of input data (abbreviations – as in table 4).

example. The value of the risk handling procedure index for the optimal solution was $RT = 2819.17$.

Figures 2b, 2c, and 2d present the "reaction" of the model to changes within its main parameters, i.e. changes in the value of the measure of risk component expressing damage/losses resulting from hazard activation.

In the case of calculations presented in figure 2b, the reaction of the optimisation model to changes significant from the point of view of risk estimation, but small in the sense of changes in the values of the model parameters. A situation involving damage to current collector (pantograph) of the tram and parts of the overhead contact line resulting from the wear of the graphite cover (contact shoe) was analysed. This was mapped in the model, increasing the value of component r_{41} to 0.500. A transfer of the symbol of this element "SO" in the cycle structure and its more frequent occurrence can be observed. The value of the risk handling procedure index for this solution was $RT = 2818.45$.

The solutions presented in figures 2c and 2d concern typically theoretical calculation cases. The threshold (minimum and maximum) values of the measure of risk component expressing damage/losses after hazard activation were used there. And so, for the calculation case in figure 2c, minimum values $r_{i1} = 0,250$ for all $(i = 1,2,...,9)$ object elements and/or all the hazards were introduced, and then the maximum values, i.e. $r_{i1} = 0,625$, for the case in figure 2d. The values of the risk handling procedure index for optimal solutions obtained in calculation cases c) and d) in figure 2 were: $RT = 2848.46$ and $RT = 2605.70$, respectively. Considerable changes in the value of the time interval between individual object maintenances should be noted.

6. Conclusions

The problem of the optimisation of the technical object maintenance system taking account of risk requires proper formulation of the objective function. It may be based on the risk model or take account of the component concerning risk. In the method developed here, an original form of the objective function was formulated. Its components are risk handling procedure costs and the values of reducing the risk measure obtained as a result of avoiding damage to object elements. In the objective function of the optimisation process, a risk model based on two criteria was applied: the criterion of damage/losses suffered as a result of hazard activation and the criterion of the possibility of these hazards' activation. Within the scope of the first criterion, a finite set of damage/loss value levels is used. The second criterion is dependent on the object's working time. Its values are expressed by the probability of the event in which the element requires maintenance activities after having worked for a specific number of working time units. It was assumed that the risk measure model is a sum of the products of risk components according to individual criteria and the significance of these criteria.

The process of making maintenance-related decisions is armed with procedures based on a risk valuation pattern. The author's original risk valuation pattern was used, classifying risk into three categories (acceptable, tolerable, and non-acceptable). When the risk measure value is included in the area of the non-acceptable or tolerable risk category, risk handling procedures (maintenance activities) should be implemented. Special cases of risk valuation resulting from the use of the patterns applied were discussed.

The structure of the maintenance cycles in preventive strategies renders the objective function of the decision-making problem discontinuous. It is therefore suggested to solve the problem with the use of dynamic programming methods. For this purpose, a mathematical notation of such modelling was developed and presented. It is a relatively difficult task to determine the range of values that the decision variables may assume in this modelling. The problem was illustrated by showing a formula for obtaining a sample acceptable solution. Moreover, a formal/mathematical notation of generating value ranges of decision variables was presented.

The complexity of the formulated optimisation task required computer assistance in order to obtain solutions. A dedicated computer program was developed and with its use, solutions to the task of optimisation of the process of planning maintenance of selected tram devices were obtained. The results of the optimisation were also prepared and their visualisation carried out with the use of the program. The program also makes it possible to check the "reaction" of the optimisation model to parameter changes. Above all, changes significant from the point of view of risk estimation, but small in the sense of changes in the values of the parameters were checked. Among others, a situation involving damage to the current collector (pantograph) of the tram and parts of the overhead contact line resulting from excessive wear of the current collector's graphite cover (contact shoe) was analysed. A significant change in the structure of the maintenance cycle and a change in the value of the *RT* index was observed. This leads us to believe that the prepared method will make it possible to solve decision-making problems concerning: the manner, scope, and schedules of replacements, repairs, and regular maintenance of technical object elements, the manner and schedules of diagnosing and preventive replacement of elements, and the problems of providing the maintenance subsystem with spare parts.

Acknowledgement

The research work financed with the means of statutory activities of Faculty of Machines and Transport Poznan University of Technology 05/52/DSPB/0259.

References

1. Amundrud Ø, Aven T. On how to understand and acknowledge risk. Reliability Engineering and System Safety 2015; 142: 42–47, http://dx.doi.org/10.1016/j.ress.2015.04.021.
2. Brennan F. Risk Based Maintenance for Offshore Wind Structures. Procedia CIRP 2013; 11: 296–300, http://dx.doi.org/10.1016/j.procir.2013.07.021.
3. Carazas F G, Souza G F M. Risk-based decision making method for maintenance policy selection of thermal power plant equipment. Energy 2010; 35: 964–975, http://dx.doi.org/10.1016/j.energy.2009.06.054.
4. Commission Implementing Regulation (EU) 2015/1136 of 13 July 2015 amending Implementing Regulation (EU) No 402/2013 on the common safety method for risk evaluation and assessment, 2015: Official Journal of the European Union.
5. Dekker R. Applications of maintenance optimization models: a review and analysis. Reliability Engineering and System Safety 1996; 51: 229–240, http://dx.doi.org/10.1016/0951-8320(95)00076-3.
6. Dickerson D E, Ackerman P J. Risk-based Maintenance Management of U.S. Public School Facilities. Procedia Engineering 2016; 145: 685–692, http://dx.doi.org/10.1016/j.proeng.2016.04.069.
7. Flage R, Aven T. Emerging risk – Conceptual definition and a relation to black swan type of events. Reliability Engineering and System Safety 2015; 144: 61–67, http://dx.doi.org/10.1016/j.ress.2015.07.008.
8. Fujiyamaa K, Nagaia S, Akikunib Y, Fujiwarab T, Furuyab K, Matsumotob S, Takagib K, Kawabatac T. Risk-based inspection and maintenance systems for steam turbines. International Journal of Pressure Vessels and Piping 2004; 81: 825–835, http://dx.doi.org/10.1016/j.ijpvp.2004.07.005.
9. Gill A. Metoda wyznaczania struktury cyklu napraw elementów pojazdów szynowych. Pojazdy Szynowe 2004; 3-4: 38-42.
10. Gill A. Procedury decyzyjne w obsługiwaniu obiektów systemów technicznych uwzględniające analizę ryzyka. [rozprawa doktorska], Poznań: Politechnika Poznańska, 2007.
11. Gill A, Kadziński A. Metoda organizacji systemu obsługiwania obiektów pojazdów szynowych uwzględniająca poziom nienaruszalności bezpieczeństwa. Organizacja transportu szynowego, in: XVII Konferencji Naukowej "Pojazdy Szynowe 2006", Kazimierz Dolny 2006: 575–585.
12. Gill A, Kadziński A. System obsługiwania pojazdów szynowych jako element w warstwowym modelu ich systemów bezpieczeństwa. Pojazdy Szynowe 2006; 4: 31-38.
13. Gill A, Kadziński A. Warunkowanie ryzykiem w procedurach decyzyjnych w obsługiwaniu pojazdów szynowych. Logistyka 2009; 6: CD.
14. Gill A, Kadziński A. The determination procedure of the onset of the object wear-out period based on monitoring of the empirical failure intensity function. Eksploatacja i Niezawodność – Maintenance and Reliability 2015; 17 (2): 282–287, http://dx.doi.org/10.17531/ein.2015.2.16.

15. Hameed A, Khan F, Ahmed S. A risk-based shutdown inspection and maintenance interval estimation considering human error. Process Safety and Environmental Protection 2016; 100: 9–21, http://dx.doi.org/10.1016/j.psep.2015.11.011.
16. Hu J, Zhang L. Risk based opportunistic maintenance model for complex mechanical systems. Expert Systems with Applications 2014; 41: 3105–3115, http://dx.doi.org/10.1016/j.eswa.2013.10.041.
17. Moradkhani A, Haghifam M R, Abedi S M. Risk-based maintenance scheduling in the presence of reward penalty scheme. Electric Power Systems Research 2015; 121: 126–133, http://dx.doi.org/10.1016/j.epsr.2014.12.006.
18. Gill A. The concept of safety system for the selected hazards identified in tram communication. Technika Transportu Szynowego 2015, 10: 2065-2074.
19. Gill A, Kadziński A. Hazard identification model. Proceedings of 20th International Scientific Conference Transport Means 2016; 5-7 Oct. Juodkrantė, Lithuania: 885-890.
20. Gill A, Kobaszyńska-Twardowska A. Identyfikacja zagrożeń w wybranych strefach tramwaju z wykorzystaniem metody Bow-Tie. Logistyka 2014; 6: 5514-5521.
21. Kadziński A. Studium wybranych aspektów niezawodności systemów oraz obiektów pojazdów szynowych. Poznań: Wyd. Politechniki Poznańskiej, Rozprawy series, no. 511, 2013.
22. Kadziński A, Gill A. Integracja pojęć. In: Krystek R, editor. Zintegrowany system bezpieczeństwa transportu. Warszawa: WKŁ, 2009: 285-288.
23. Kadziński A, Warguła J, Gill A, Szacowanie i wartościowanie ryzyka zagrożeń związanych z odcinkiem szybkiego tramwaju na poznańskiej sieci tramwajowej. Logistyka 2012; 3: 939-948.
24. Kadziński A, Woźniak A. O modelach kwantyfikacji zagrożeń bezpieczeństwa w transporcie kolejowym, in: Mat. XIV konferencji naukowej "Pojazdy szynowe"; Kraków-Arłamów, 2000: 129-136.
25. Khan F I, Haddara M R. Risk-based maintenance of ethylene oxide production facilities. Journal of Hazardous Materials 2004; 108(3): 147–159, http://dx.doi.org/10.1016/j.jhazmat.2004.01.011.
26. Kosmowski K T. Metodyka analizy ryzyka w zarządzaniu niezawodnością i bezpieczeństwem elektrowni jądrowych. Gdańsk: Wydawnictwo Politechniki Gdańskiej, 2003.
27. Młyńczak M, Pisarski P. Wspomaganie komputerowe w prognozowaniu obsług obiektów mechanicznych: Metody prognozowania w inżynierii niezawodności, in: Mat. XXXI Zimowej Szkoły Niezawodności; Szczyrk, 2003: 356-364.
28. Nilsson F. Risk based approach to plant life management. Nuclear Engineering and Design 2003; 221(1-3): 293-300, DOI: 10.1016/S0029-5493(02)00334-5.
29. Nielsen J J, Sørensen J D. On risk-based operation and maintenance of offshore wind turbine components. Reliability Engineering and System Safety 2011; 96: 218–229, http://dx.doi.org/10.1016/j.ress.2010.07.007.
30. Niu G, Yang B, Pecht M. Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance. Reliability Engineering and System Safety 2010; 95: 786–796, http://dx.doi.org/10.1016/j.ress.2010.02.016.
31. Niziński S. Eksploatacja obiektów technicznych. Radom: Wydawnictwo i Zakład Poligrafii Instytutu Technologii Eksploatacji, 2002.
32. Ochodek B, Ochodek M. Algorytmy i struktury danych. Piła: Państwowa Wyższa Szkoła Zawodowa w Pile, 2003.
33. Okulewicz J, Salamonowicz T. Porównanie wybranych strategii odnów profilaktycznych. In: Mat. XXXIV Zimowej Szkoły Niezawodności, 2006: 218-227.
34. Ponchet A, Fouladirad M, Grall A. Assessment of a maintenance model for a multi-deteriorating mode system. Reliability Engineering and System Safety 2010; 95: 1244–1254, http://dx.doi.org/10.1016/j.ress.2010.06.021.
35. Rusin A, Wojaczek A. Optimization of power machines maintenance intervals taking the risk into consideration. Eksploatacja i Niezawodność – Maintenance and Reliability 2012; 14 (1): 72-76.
36. Sarkar A, Panja S H, Sarkar B. Survey of maintenance policies for the Last 50 Years. International Journal of Software Engineering & Applications (IJSEA) 2011; 2(3)130-148.
37. Straub D, Faber M H. Risk based inspection planning for structural systems. Structural Safety 2005; 27: 335–355, http://dx.doi.org/10.1016/j.strusafe.2005.04.001.
38. Tang D, Yu J, Chen X, Makis V. An optimal condition-based maintenance policy for a degrading system subject to the competing risks of soft and hard failure. Computers & Industrial Engineering 2015; 83: 100–110, http://dx.doi.org/10.1016/j.cie.2015.02.003.
39. Wang L, Chu J, Mao W. A condition-based order-replacement policy for a single-unit system. Applied Mathematical Modelling 2008; 32: 2274–2289, http://dx.doi.org/10.1016/j.apm.2007.07.016.
40. Zio E. Reliability engineering: Old problems and new challenges, Reliability Engineering & System Safety 2009; 94(2): 125–141, http://dx.doi.org/10.1016/j.ress.2008.06.002.

Adrian GILL
Institute of Combustion Engines and Transport
Poznan University of Technology
Piotrowo 3, 60-965 Poznan, Poland

E-mail: adrian.gill@put.poznan.pl

Article citation info:
SUN L, GU X-H, SONG P, DI Y. A generalized equivalent temperature model in a time-varying environment. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 432–440, http://dx.doi.org/10.17531/ein.2017.3.14.

Li SUN
Xiao-Hui GU
Pu SONG
Yi DI

A GENERALIZED EQUIVALENT TEMPERATURE MODEL IN A TIME-VARYING ENVIRONMENT

UOGÓLNIONY MODEL TEMPERATURY RÓWNOWAŻNEJ W ZMIENNYM W CZASIE OTOCZENIU

Accelerated degradation test at high temperature level is a common method to accelerate the degradation of products by elevating temperature, and the obtained degradation data are then used to obtain the estimate of the performance at normal temperature after extrapolating the degradation through accelerating model. However, the normal temperature is ever-changing rather than constant. Therefore, a generalized equivalent temperature model based on power law degradation path is proposed to establish a connection between accelerated degradation data and degradation data at normal temperature. The model takes the equal degradation measure as a principle and the conclusion is demonstrated that the increments of the degradation under the same magnitude, same time and different orders of temperature stresses are same. The result shows that the empirical equivalent temperature model is a special case of the proposed model. The accuracy of the proposed model is finally demonstrated by a case study of nitrile rubber O-rings.

Keywords: *accelerated degradation test, equivalent temperature, nature storage, degradation path, rubber O-rings.*

Przyspieszone badania degradacji (badania starzeniowe) prowadzone w warunkach wysokiej temperatury stanowią powszechnie stosowaną metodę przyspieszania starzenia produktów poprzez podwyższanie temperatury. Otrzymane w takich badaniach dane degradacyjne wykorzystuje się do szacowania wydajności produktu w temperaturze normalnej na zasadzie ekstrapolacji. Głównym ograniczeniem tej metody jest fakt, że normalna temperatura nie jest stała lecz zmienia się w czasie. Dlatego też, aby skorelować dane z przyspieszonej degradacji z danymi dotyczącymi starzenia w normalnej temperaturze, zaproponowaliśmy uogólniony model temperatury równoważnej oparty na krzywej degradacji opisanej prawem potęgowym. W modelu przyjęto zasadę równego stopnia degradacji i wykazano, że przyrosty degradacji przy tej samej wartości i czasie działania naprężeń termicznych różnego rzędu są takie same. Wyniki pokazują, że empiryczny model temperatury równoważnej jest szczególnym przypadkiem proponowanego przez nas modelu. Trafność opisanego w pracy modelu wykazano na podstawie studium przypadku dotyczącego uszczelek nitrylowych, tzw. oringów.

Słowa kluczowe: *przyspieszone badania starzenia, temperatura równoważna, przechowywanie w warunkach naturalnych, krzywa degradacji, oringi nitrylowe.*

1. Introduction

Accelerated test is an efficient method to collect information of products by measuring performance data directly over time from the test at high stress while the collected data are used to extrapolate the information through a physically reasonable statistical model to obtain the estimate of life or long-term performance at lower stress, normal use or storage condition [12]. The frequently used stresses include use rate, voltage, humidity, pressure, especially temperature [13, 16, 23]. Long-term reliability of gold (Au) and copper (Cu) ball bonds in fineline ball grid array package under storage condition 30°C was estimated by high temperature storage bake test at elevated temperatures of 150°C, 175°C and 200°C [8]. Wang predicted the storage life of aerospace electromagnetic relay under storage temperature 25°C -32°C based on auto-regressive and moving average model and wavelet transform model [21]. Anisotropic magnetoresistive read sensors were exposed to elevated temperatures to estimate end-of-life conditions under normal operating temperatures [5]. Huang predicted the life of tantalum capacitors under working temperature which was specified as 35°C [4]. Vakulov studied the properties changing in storage condition by accelerating ageing test on the example of rubbers K-14-1 and K-14-2 [18].

From the examples above, a common phenomenon is observed that the temperature in storage condition or normal operating is often assumed as a constant temperature [7](e.g. 25°C,30°C), or a temperature interval. However, the temperature in real storage or operating condition often varies with the season and region. For some long-life products, a minor temperature difference may lead to a major difference in the result of life assessment. As a result, it is important to consider the impact of the changing temperature appropriately. Particularly, it is of interest to wonder whether there is a temperature under which the life or performance of the product is equal to the life or performance at real storage or operating temperature. This issue consists of the primary goal of this paper, proposing an equivalent temperature model.

In literature, a substantial number of degradation models have been developed to model the accelerated degradation data [10]. In general, these existing models can be divided into linear and nonlinear

models. In these models, power law degradation path model can describe both linear degradation path and nonlinear degradation path by different powers, and several empirical studies show that the expected degradation at time t is often proportional to a power law [2, 17, 19]. Moreover, this kind of model is widely used in engineering fields like electrical connectors, rubber materials [9], green laser diodes [11] and film resistances [1]. In addition, the power law degradation path has also been used to optimize the accelerated degradation experiment. These observations indicate the wide application of the power law model [24, 25]. Accordingly, the proposed equivalent temperature model is based on power law degradation path due to its flexibility and good performance.

In the inference process of the model, the conclusion is demonstrated to simplify the model that the increments of the degradation under the same magnitude, the same time and different orders of temperature stresses are same. Meanwhile, the influences to equivalent temperature of the two key parameters in the model, the time index α and accelerated parameter β, are analyzed by taking the first partial derivative of the equivalent temperature model. The result shows that equivalent temperature would decrease with the increasing of β and α. Besides, the model is compared with the average temperature model and the empirical equivalent temperature model which is a special case of the proposed model in the linear case while α=1. Thus the proposed model is also known as generalized equivalent temperature model. Moreover, the validity of the proposed model is proved by rubber O-rings data from the accelerated degradation test and the nature storage test. Finally, the fitness of three equivalent temperature models are compared by the criterion of mean squared error (MSE).

Compared with accelerated degradation test, degradation test under normal temperature could obtain more precise results but become increasingly difficult owing to the characteristics of long-period and high-cost which conflicts with the marketplace demands for decreasing development time. In this case, equivalent temperature model establishes an effective connection between accelerated degradation test and degradation test at normal temperature. In terms of actual usage, the equivalent temperature of nature storage test can be obtained by directly substituting the statistical data of temperatures in normal condition and the relevant parameters derived by accelerated degradation test into the model. By means of this model, the life or degradation performance information of the product under normal non-constant temperature can be predicted directly without the time-consuming and costly degradation test. The model can be applied to all of the products whose degradation follows the power law degradation path.

The remaining segments of this paper are organized as follows. Section 2 introduces the derivation of the empirical equivalent temperature model and the generalized equivalent temperature model. In Section 3, parameters in the model are estimated based on accelerated degradation data. Section 4 introduces the sensitivity analysis of parameters to equivalent temperature. A case study of rubber O-rings is presented to verify the high performance of generalized equivalent temperature model compared with other two equivalent temperature models in Section 5. Section 6 presents the conclusion of the paper and application of the model.

2. Equivalent temperature model

2.1. Empirical equivalent temperature model

The relationship between the degradation rate and temperature T could be descripted as Arrhenius function which is written as [14, 20]:

$$\frac{dy}{dt} = Ae^{\left[\frac{-E}{RT}\right]} \tag{1}$$

where y is the degradation, dy/dt is the degradation rate, A is the pre-exponential factor of the Arrhenius rate constant, E represents material apparent activation energy, R is the universal gas constant(8.3145 kJ kmol^{-1} K^{-1}) and T is the absolute temperature in Kelvin. Specify ω=ln$A$, β=-E/R to get the simple model of the function and it is obviously that β<0 from its physics meaning. Then the degradation can be derived as:

$$y = e^{[\omega+\beta/T]}t + y_0 \tag{2}$$

where y_0 is a constant.

Accelerated degradation test accelerates the degradation of products by means of increasing the temperature stress to obtain the degradation at accelerated temperature and the temperature stress is assumed constant during the test [3]. However, the temperature stress is often changing during nature storage test which is recognized that $T=f(t)$. It is assumed that there is an equivalent temperature T_0 during which the degradation in a certain time ζ is equal to the degradation at a time-varying temperature $T(t)$ in the same time. As such, the following equation must be satisfied:

$$e^{[\beta/T_0]}\varsigma = \int_0^{\varsigma} e^{[\beta/T(t)]}dt\ . \tag{3}$$

For a giving $T(t)$, the equivalent temperature T_0 can be simplified as:

$$T_0 = \frac{\beta}{\ln\left\{\frac{1}{\varsigma}\int_0^{\varsigma} e^{[-\beta/(T(t))]}dt\right\}} \tag{4}$$

However, in real environment, the equation of T_0 is difficult to have its analytical expression because $T(t)$ is unavailable to model continuously. So the integral expression in denominator is discretized, then:

$$\mathbf{M_1}: T_0 = \frac{\beta}{\ln\left\{\frac{1}{\varsigma}\sum_{i=1}^{n} t_i \cdot \exp\left(\frac{\beta}{T_i}\right)\right\}} \tag{5}$$

where $t_i(i=1,2,...,n)$ is the time at temperature T_i and the total storage time is $\varsigma = \sum_{i=1}^{n} t_i$.

2.2. Generalized equivalent temperature model

However, sometimes the degradation rate k is correlative with test time t, then Eq. (1) and Eq. (2) can be rewritten as:

$$\begin{cases} \dfrac{dy}{dt} = \dfrac{1}{\alpha} e^{[\omega+\beta/T]} t^{\alpha-1} \\ y = e^{[\omega+\beta/T]} t^{\alpha} + y_0 \end{cases} \tag{6}$$

This form of degradation process is also called power law model where time index α is a constant related to the degradation mechanism with α>0. The degradation path is an up-convex curves while 0<α<1, and a straight line while α=1, and an up-concave curves while α>1.

It is also assumed that y_0=0 without loss of generality by certain corresponding transformation.

Similarly, in time-varying environment, the equivalent temperature T_0 can be expressed as:

$$T_0 = \frac{\beta}{\ln\left\{\frac{1}{\alpha\varsigma^{\alpha}}\int_0^{\varsigma} e^{\left[\beta/T(t)\right]} t^{\alpha-1} dt\right\}} \tag{7}$$

For equation (4) and (7), the expression of T_0 in equation (7) can be reduced to the equation (4) exactly while α=1. Undoubtedly, the analytical expression of T_0 in equation (7) is hard to handle even by discretizing. Based on the equal degradation measure principle, a generalized equivalent temperature model is derived.

For the requirement of derivation of the generalized equivalent temperature model, the following assumption is introduced.

The increments of the degradation under same magnitude, same time and different orders of temperature stresses are same in natural environment.

To illustrate the idea of the above assumption, Fig. 1 is given. As shown in Fig.1, it is assumed that a product is aged at the temperature T_1 for time t_1 first and the increment of the degradation is y^{11}. Then, the product is aged at temperature T_2 for time t_2, and the increment of the degradation is y^{12}. The degradation path is shown as the short dash line in Fig. 1. In addition, the degradation rate is k_1 at the temperature T_1, and k_2 at the temperature T_2. In general, t_1, t_2, k_1, k_2 are greater than 0.

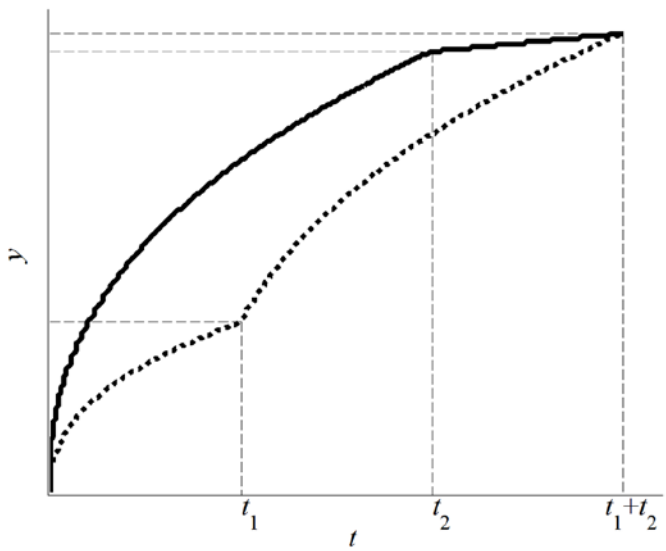


Fig. 1. The degradation path under different stress orders

From Eq. (6), we can get:

$$k_1 = \exp\left(\omega + \beta / T_1\right) \tag{8}$$

$$k_2 = \exp\left(\omega + \beta / T_2\right). \tag{9}$$

$$y^{11} = k_1 t_1^{\alpha}. \tag{10}$$

The degradation is continuous and only depends on the accumulated degradation and the current temperature stress. As a result, the degradation y^{11} should be converted to the equivalent time τ_{12} at temperature T_2:

$$y^{11} = k_2 \tau_{12}^{\alpha}. \tag{11}$$

Put Eq.(8), Eq.(9) and Eq.(10) into Eq.(11) and we can obtain:

$$\tau_{12} = \left(\frac{k_1}{k_2}\right)^{1/\alpha} = \left\{\exp\left[\beta\left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right]\right\}^{1/\alpha}. \tag{12}$$

Through the derivation above, y^{12} can be formulated as:

$$y^{12} = k_2\left(t_2 + t_1\left(\frac{k_1}{k_2}\right)^{1/\alpha}\right)^{\alpha}.$$

On the contrary, the product is aged at the temperature T_2 for time t_2 firstly and the increment of the degradation is y^{21}. Then, the product is aged at the temperature T_1 for time t_1, and the increment of the degradation is y^{22}. The degradation path is shown as the heavy line in Fig. 1. Similarly, we have:

$$y^{21} = k_2 t_2^{\alpha}$$

$$y^{22} = k_1\left(t_1 + t_2\left(\frac{k_2}{k_1}\right)^{1/\alpha}\right)^{\alpha}$$

It is proved that y^{12}=y^{22} and the proof is given as follows:

$$\frac{y^{12}}{y^{22}} = \frac{k_2\left(t_2 + t_1\left(\frac{k_1}{k_2}\right)^{1/\alpha}\right)^{\alpha}}{k_1\left(t_1 + t_2\left(\frac{k_2}{k_1}\right)^{1/\alpha}\right)^{\alpha}} = \frac{\left((k_2)^{1/\alpha} t_2 + t_1 (k_1)^{1/\alpha}\right)^{\alpha}}{\left((k_1)^{1/\alpha} t_1 + t_2 (k_2)^{1/\alpha}\right)^{\alpha}} = 1$$

In the nature storage test, the storage time is t_i(i=1,2,...,n) at temperature T_i and the total storage time is $\varsigma = \sum_{i=1}^{n} t_i$. The degradation after the whole nature storage test is equivalent to the degradation at temperature T_0 for time ζ. The generalized equivalent temperature model can be written as:

$$\mathbf{M_0}: \begin{cases} T_0 = \dfrac{1}{\dfrac{1}{T_n} + \dfrac{\alpha}{\beta}\ln(\tau_n) - \dfrac{\alpha}{\beta}\ln(\varsigma)} \\ \tau_n = t_n + (\tau_{n-1})\left(\dfrac{k_{n-1}}{k_n}\right)^{1/\alpha} \\ \tau_{n-1} = t_{n-1} + \tau_{n-2}\left(\dfrac{k_{n-2}}{k_{n-1}}\right)^{1/\alpha} \\ \cdots \\ \tau_2 = t_2 + \tau_1\left(\dfrac{k_1}{k_2}\right)^{1/\alpha} \\ \tau_1 = t_1 \end{cases} \tag{13}$$

where $\tau_1, \tau_2, ..., \tau_n$ are equivalent time, and $\frac{k_{n-1}}{k_n} = \exp\left(\beta\left(\frac{1}{T_{n-1}} - \frac{1}{T_n}\right)\right)$.

3. Parameter estimation based on accelerated degradation data

In constant accelerated degradation test, there are N accelerated temperature stresses, and the number of the samples under each stress are n_i. The degradation path is described by Eq (14).

$$y_{ijl}\left(t_{ijl} \mid T_i\right) = y_{ij0} + k_{ij} t_{ijl}^{\alpha}, \quad i = 1,2,\cdots,N, j = 1,2,\cdots,n_i, l = 1,2,\cdots,L_{ij} \tag{14}$$

where $y_{ijl}(t_l|T_i)$ is the measured degradation of the jth sample at time t_l under temperature T_i, y_{ij0} is the initial value of the degradation and it is assumed that y_{ij0}=0 without loss of generality, k_{ij} is the degradation rate of the jth sample at temperature T_i and k_{ij}>0, L_{ij} is the mumber of measurement time for each sample.

First, the value of α is calculated by a least square method which minimize the value of I and the expression of I is shown as the formula below

$$\begin{cases} I(\alpha) = \sum_{i=1}^{N}\sum_{j=1}^{n_i}\sum_{l=1}^{L_{ij}}\left(y_{ijl} - \hat{y}_{ijl}\right)^2 \\ \hat{y}_{ijl} = \hat{k}_{ij} t_l^{\alpha} \end{cases} \tag{15}$$

where the value of $\hat{k}_{ij}$ can be estimated by linear fit with an exact value of α and $\hat{y}_{ijl}$ is the estimated degradation.

Then, combing Eq.(6) and Eq.(15), the following relationship can be obtained

$$k_{ij}\left(T_i\right) = e^{\left[\omega + \beta / T_i\right]} \tag{16}$$

Taking the logarithm of the left and right sides of the above formula, the parameter ω, β can be estimated by a least square method.

4. Sensitivity analysis of parameters

The equivalent temperature in M_0 is closely related to the parameters β and α. Next, we would discuss how equivalent temperature varies with them.

4.1. Sensitivity analysis of parameter β

Taking the first partial derivative of Eq. (13) with respect to β, the following equations can be given.

$$\begin{cases} \dfrac{\partial T_0}{\partial \beta} = -\left(\dfrac{1}{\dfrac{1}{T_n} + \dfrac{\alpha}{\beta}\ln(\tau_n) - \dfrac{\alpha}{\beta}\ln(\varsigma)}\right)^2 \cdot \left[\dfrac{\alpha}{(\beta)^2}\left(\ln(\varsigma) - \ln(\tau_n)\right) + \dfrac{\alpha}{\beta\tau_n}\dfrac{d\tau_n}{d\beta}\right] \\ \dfrac{d\tau_n}{d\beta} = \dfrac{d\tau_{n-1}}{d\beta}\left(\dfrac{k_{n-1}}{k_n}\right)^{1/\alpha} + \tau_{n-1}\cdot\dfrac{1}{\alpha}\left(\dfrac{k_{n-1}}{k_n}\right)^{\frac{1}{\alpha}}\cdot\left(\dfrac{1}{T_{n-1}} - \dfrac{1}{T_n}\right) \\ \cdots \\ \dfrac{d\tau_2}{d\beta} = \dfrac{d\tau_1}{d\beta}\left(\dfrac{k_1}{k_2}\right)^{1/\alpha} + \tau_1\cdot\dfrac{1}{\alpha}\left(\dfrac{k_1}{k_2}\right)^{\frac{1}{\alpha}}\left(\dfrac{1}{T_1} - \dfrac{1}{T_2}\right) \\ \dfrac{d\tau_1}{d\beta} = 0 \end{cases} \tag{17}$$

It could be obtained that $\lim_{\beta\to-\infty}\frac{\partial T_0}{\partial\beta} = 0$, $\lim_{\alpha\to\infty}\frac{\partial T_0}{\partial\beta} = 0$ by Eq. (17). Because of the complexity of $\frac{\partial T_0}{\partial\beta}$, it is achieved by numerical calculation rather than analytical approaches. Take a storehouse in Nanjing for example. Its daily average temperature in 2014 is shown as Fig.2. The partial derivative of equivalent temperature to β is shown in Fig.3 in which the range of β is -10000~0 and the range of α is 0~5. The partial derivative of equivalent temperature to β is less than 0 in the numerical range of Fig.4 and it increases with the rise of α but decreases with the rise of β. Combined with the limit value, we can get the conclusion that equivalent temperature would decreases with the increase of β but its descent rate would gradually increase while the value of α is generally taken as a constant.

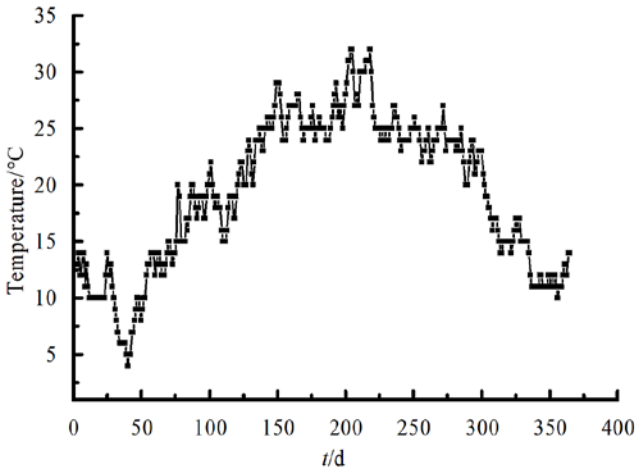


Fig. 2. Daily average temperature in 2014 of a storehouse in Nanjing

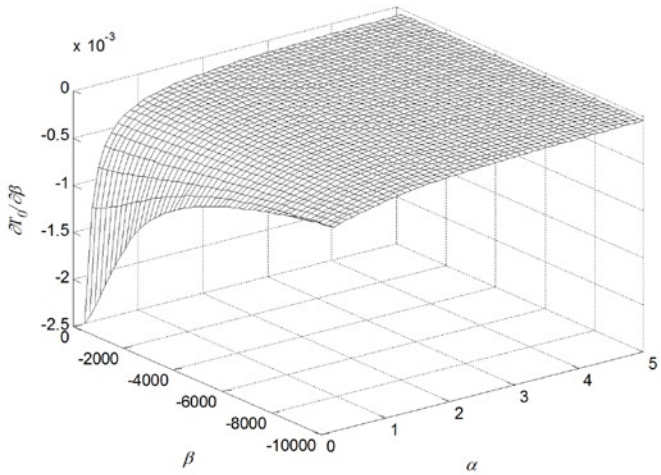


Fig. 3. The partial derivative of equivalent temperature to β

4.2. Sensitivity analysis of parameter α

Taking the first partial derivative of Eq. (13) with respect to α, it can be obtained

$$
\begin{cases}
\dfrac{\partial T_0}{\partial \alpha} = -\left(\dfrac{1}{\dfrac{1}{T_n}+\dfrac{\alpha}{\beta}\ln(\tau_n)-\dfrac{\alpha}{\beta}\ln(\varsigma)}\right)^2 \cdot \left[\dfrac{1}{\beta}\left(\ln(\tau_n)-\ln(\varsigma)\right)+\dfrac{\alpha}{\beta\tau_n}\dfrac{d\tau_n}{d\alpha}\right] \\
\dfrac{d\tau_n}{d\alpha} = \dfrac{d\tau_{n-1}}{d\alpha}\left(\dfrac{k_{n-1}}{k_n}\right)^{1/\alpha} + (\tau_{n-1})\left(\dfrac{k_{n-1}}{k_n}\right)^{1/\alpha}\ln\left(\dfrac{k_{n-1}}{k_n}\right)\left(-\dfrac{1}{\alpha^2}\right) \\
\dfrac{d\tau_{n-1}}{d\alpha} = \dfrac{d\tau_{n-2}}{d\alpha}\left(\dfrac{k_{n-2}}{k_{n-1}}\right)^{1/\alpha} + (\tau_{n-2})\left(\dfrac{k_{n-2}}{k_{n-1}}\right)^{1/\alpha}\ln\left(\dfrac{k_{n-2}}{k_{n-1}}\right)\left(-\dfrac{1}{\alpha^2}\right) \\
\cdots \\
\dfrac{d\tau_2}{d\alpha} = \dfrac{d\tau_1}{d\alpha}\left(\dfrac{k_1}{k_2}\right)^{1/\alpha} + \tau_1\left(\dfrac{k_1}{k_2}\right)^{1/\alpha}\ln\left(\dfrac{k_1}{k_2}\right)\left(-\dfrac{1}{\alpha^2}\right) \\
\dfrac{d\tau_1}{d\alpha} = 0
\end{cases}
\tag{18}
$$

where $\ln\left(\dfrac{k_{n-1}}{k_n}\right)=\beta\left(\dfrac{1}{T_{n-1}}-\dfrac{1}{T_n}\right)$.

Similarly, the partial derivative of equivalent temperature to α is shown in Fig.5 in which the range of β is -10000~0 and the range of α is 0~5. The partial derivative of equivalent temperature to α is less than 0 in the numerical range of Fig.4 and it increases with the rise of both α and β. So we can get the conclusion that equivalent temperature would decreases with the increase of α but its descent rate would gradually reduce while the value of β is generally taken as a constant.

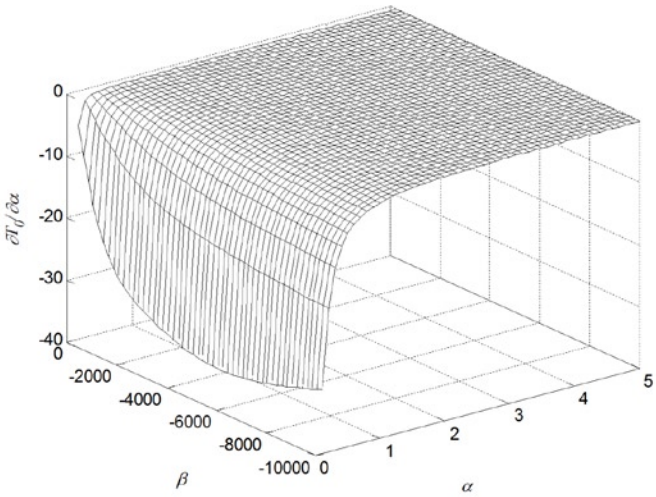


Fig. 4. The partial derivative of equivalent temperature to α

5. Practical Calculation

5.1. Experiment

5.1.1. Accelerated degradation test

Accelerated degradation test of nitrile rubber O-rings was carried out at 50°C, 60°C, 70°C and 80°C, with 8 samples at each temperature [22]. In order to simulate the stress state of nitrile rubber O-rings in practical application environment, the sample was placed in a specially-made jig which was composed of two compression plates and a spacer ring to control the compression ratio, as shown in Fig. 5. Tight the four bolts to compress the rubber O-rings to achieve the specified thickness while putting the rubber O-rings in the jigs.

When rubber O-rings are held under compression, physical or chemical changes could occur that prevent the rubber returning to its original dimensions after the release of the deforming force. The result is a compression set and the magnitude of which depends on the time and temperature of compression as well as on the time and temperature of recovery. At elevated temperature, chemical changes become increasing more important and lead to a permanent [6]. As such, the compression set ε was chosen as the performance parameter, which is formulated as:

$$\varepsilon = \frac{D_0 - D_t}{D_0 - D_x} \times 100\%$$

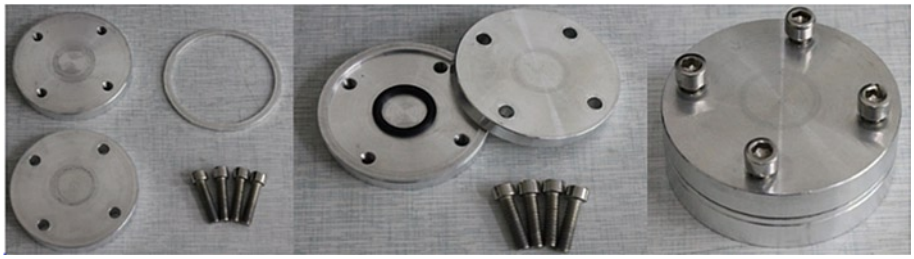


Fig. 5. The specially-made jig

where D_0 is the initial thickness of the test sample, D_x is the height of the spacer ring, D_t is the sickness after time t in the aging chamber which is measured by releasing the deforming force on the samples and allowing a 1 hour recovery period at 25°C. ε=31% is set as the failure threshold for rubber O-rings since the internal air tightness of the structure will be destroyed along with the reduction of safety and reliability while ε is greater than 31%.

Place the loaded rubber O-rings with jigs to aging chamber, after the required test duration, remove the compression assembly from the aging chamber, allow it to cool to a standard laboratory temperature 25°C for 1 h, then release the O-rings, and after a further 1 h at standard laboratory temperature, measure their thickness. Each O-ring was measured at four points around the circumference and the results were averaged. Put the measured thickness at each time point to Eq. (12) and get the compression set curve as shown in Fig. 6.

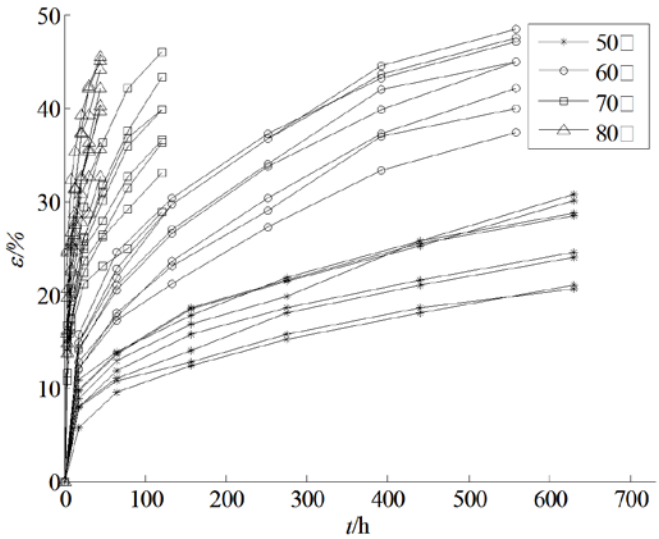


Fig. 6. The compression set curve of accelerated degradation test

5.1.2. Nature storage test

The nature storage test of 8 nitrile rubber O-rings was carried out at laboratory in June 2013. The rubber O-rings were preserved at room temperature and prevented from light, refrigerating equipment in summer and heating equipment in winter. The jigs, samples, installation method and measurement method were same with the accelerated degradation test. The daily average temperatures during the test are shown in Fig. 7. By the end of the last measurement, the test was carried out for 856 days and there were 12 data were measured for each sample. Fig.8 shows the connection of compression set between adjacent measuring points. Non-monotonicity is observed from Fig. 8 compared with the data from accelerated degradation test.

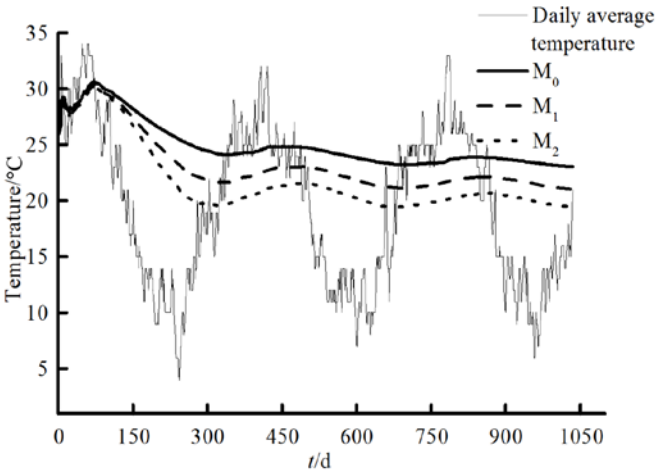


Fig. 7. Comparison of equivalent temperature under M_0, M_1 and M_2 during the nature storage test

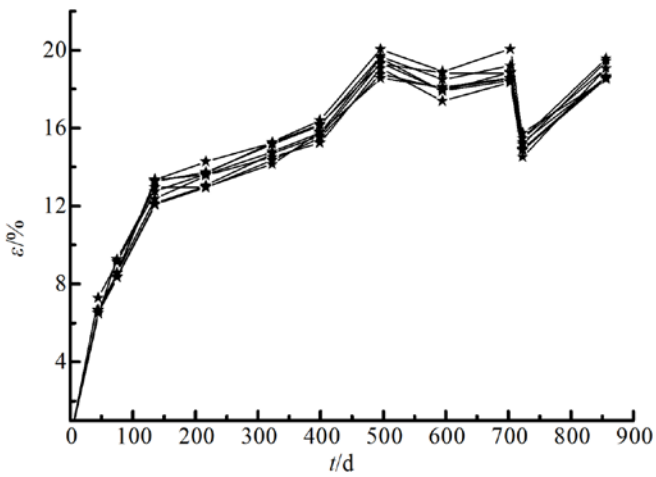


Fig. 8. The compression set curve of nature storage test

5.2. Parameter estimation of accelerated degradation test

The degradation of nitrile rubber O-rings is a function of compression set which can be formulated as:

$$y = -\ln(1-\varepsilon). \tag{19}$$

The data obtained from accelerated degradation test can be fitted by Eq.(14). The estimate of the parameter α is 0.37 and the results of the parameter k_{ij} are shown as Table 1.

Table 1. Degradation rate $\hat{k}_{ij}$ under accelerated degradation test

Stress number i	Temperature stress T_i	Sample serial number (j)	Degradation rate $\hat{k}_{ij}$
1	323.15	1	0.0315
		2	0.0310
		3	0.0255
		4	0.0305
		5	0.0296
		6	0.0244
		7	0.0207
		8	0.0213
2	333.15	1	0.0537
		2	0.0594
		3	0.0598
		4	0.0484
		5	0.0604
		6	0.0551
		7	0.0465
		8	0.0424
3	343.15	1	0.0858
		2	0.0922
		3	0.0795
		4	0.0889
		5	0.1053
		6	0.0757
		7	0.0700
		8	0.0610
4	353.15	1	0.1280
		2	0.1455
		3	0.1415
		4	0.1249
		5	0.1567
		6	0.1242
		7	0.1148
		8	0.1007

Then the value of ω and β are estimated by a least square method as Table 2 and the correlation coefficient r is greater than critical value at 1% level which validates the good fit to Arrhenius function.

$$\ln k_i = 14.7218 - 5913.0224 / T_i$$

Table 2. The parameters and correlation coefficients of accelerated model

Distribution parameters	ω	β	r
	14.7218	-5913.0224	0.9665

5.3. Model comparison

For comparative studies, we refer to the model M_1 for comparison which is expressed as Eq. (15). In addition we choose the average temperature model as M_2 which is shown as Eq. (20). And the proposed model is written as M_0.

$$\mathbf{M_2}: T_0 = \frac{1}{\varsigma}\sum_{i=1}^{n} t_i T_i \tag{20}$$

Based on M_0, M_1, M_2, we respectively obtain the equivalent temperatures of these models as shown in Table 4. The 2.36°C temperature difference between M_0 and M_1 and the 3.84°C temperature difference between M_0 and M_2 are observed from Table 3.

Table 3. Comparison of T_0 under M_0, M_1, M_2

	M_0	M_1	M_2
T_0/°C	23.32	20.96	19.48

5.3.1. The effect of β on equivalent temperature for the three models

The effect of parameter β on equivalent temperature is shown in Fig.9 that the range of β is -100000~0. From Fig. 9, we can see that both T_0 from M_0 and M_1 both increase with the decrease of β. At the same time, their increase rates decrease gradually which is consistent with our discussion in 4.1.

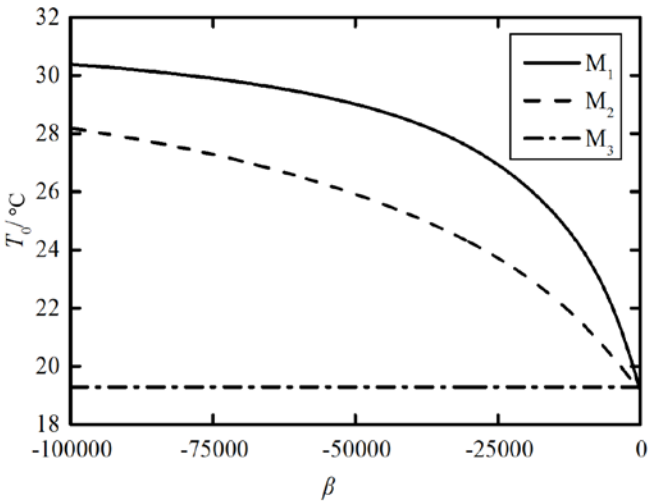


Fig. 9. The effect of parameter β on T_0

5.3.2. The effect of α to equivalent temperature for the three models

The effect of parameter α on equivalent temperature is shown in Fig.10 that the range of α is 0~10. From Fig. 10, we can see that T_0 decreases with the increase of α in M_0. At the same time, its decrease rate decrease gradually which is consistent with our discussion in 4.2. Moreover, the equivalent temperatures in M_0 and M_1 are equal while α=1 which indicates that the proposed model is more general and M_1 is a special case of the proposed model in the linear case.

5.3.3. Model verification

Substitute the daily average temperature in Fig. 8 to M_0, M_1 and M_2, respectively. The equivalent temperatures of these three models

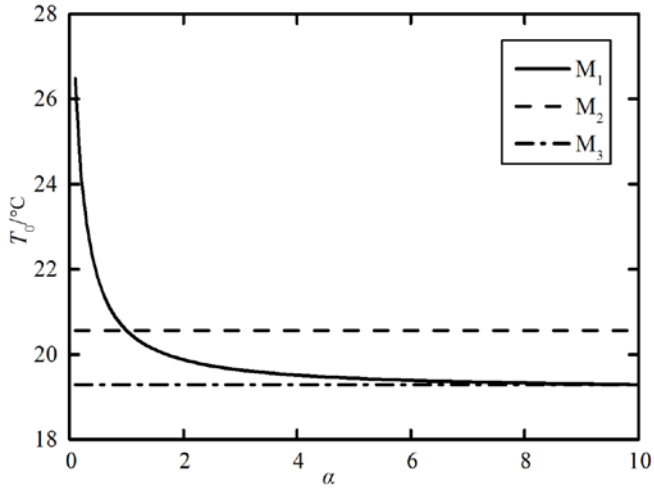


Fig. 10. The effect of parameter α on T_0

are shown in Fig.7 during the nature storage test. From Fig.7, we can see that the change range of equivalent temperature along with the change of daily average temperature is relatively large in the first year of nature storage test. However, the equivalent temperature will not produce significant motion following daily average temperature during the second year of the test. Accordingly, there is a more slight fluctuation of equivalent temperature in the third year compared with the second year. And it would be predicted that equivalent temperature will eventually approach to a constant temperature while testing time is long enough.

Degradation rates during the nature storage test can be calculated by substituting equivalent temperature in Fig.7 to accelerated model shown as Eq.(6). To quantitatively compare the above models, the MSE is used to assess the estimation accuracy of the model. Let $\hat{y}_j(t_l)$ demote the estimated value of degradation of rubber O-rings at time t_l for the jth sample, and $y_j(t_l)$ can be calculated by the measured compression set in Fig.8.

$$MSE = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{L_j}\sum_{l=1}^{L_j}\left(\hat{y}_j(t_l) - y_j(t_l)\right)^2 .$$

Using the MSE, the smallest MSE value corresponds to the best estimation accuracy [15], and thus it could serve as the criteria for model selection. For comparison, the MSEs under M_0, M_1 and M_2 is listed in Table 4. From Table 4, it we can be observed that MSE of M_0 is the smallest and the result of M_0 has higher precision in estimating the degradation of rubber O-rings in nature storage condition.

Table 4. MSEs under M_0, M_1, M_2

	M_0	M_1	M_2
MSE	0.0005676	0.0008791	0.0015947

6. Conclusion

Motivated by the need of practical engineering, a generalized equivalent temperature model is proposed in this paper. Therefore, the degradation at normal temperature which is ever-changing rather than constant can be derived by accelerated degradation test. The degradation at changing temperature for quite a time is equal

to the degradation at equivalent temperature for the same time. And the influences of the two key parameters in the model are analyzed by taking the first partial derivative of the equivalent temperature model.

It is demonstrated that the empirical equivalent temperature model is a special case of the proposed model by numerical computation. Meanwhile, the excellent performance of the proposed model is verified by nitrile rubber O-rings nature storage test. The result shows that the proposed model has the minimum MSE compared with other two equivalent temperature models.

The proposed model can be widely used in other products whose degradation at time t is proportional to a power law degradation model. For the long-lifetime products, the parameter α and parameter β in the model can be calculated by accelerated temperature stress test firstly. Then extrapolate the information in accelerated condition to equivalent temperature condition to obtain the information of life or degradation parameters in nature storage condition or real-use condition directly. It is quite time-saving and could get results with good accuracy.

References

1. Chan C K, Boulanger M, Tortorella M. Analysis of parameter-degradation data using life-data analysis programs. Proceedings Annual Reliability and Maintainability Symposium 1994; 288-291, http://dx.doi.org/10.1109/RAMS.1994.291122.
2. Cui Z, Liou J J, Yue Y. A new extrapolation method for long-term degradation prediction of deep-submicron MOSFETs. IEEE Transactions on Electron Devices 2003; 50(5): 1398-1401, http://dx.doi.org/10.1109/TED.2003.813473.
3. Escobar L A, Meeker W Q. A review of accelerated test models. Statistical science 2006; 552-577, http://dx.doi.org/10.1214/088342306000000321.
4. Huang J, Gao C, Cui W, et al. Lifetime prediction for tantalum capacitors with multiple degradation measures and particle swarm optimization based grey model. Journal of Central South University 2012; 19: 1302-1310, http://dx.doi.org/10.1007/s11771-012-1142-y.
5. Iben I E T. Head reliability of AMR sensors based on thermal stress tests. IBM Journal of Research and Development 2003; 47(4): 415-428, http://dx.doi.org/10.1147/rd.474.0415.
6. ISO/PRF Rubber, vulcanized or thermoplastic Determination of compression set Part At ambient or elevated temperatures. ISO/PRF 815-1 - Rubber, vulcanized or thermoplastic -- Determination of compression set -- Part 1: At ambient or elevated temperatures. ISO CT 61 - Plastics.
7. Lee H M, Wu J W, Lei C L. Assessing the lifetime performance index of exponential products with step-stress accelerated life-testing data. IEEE Transactions on Reliability 2013; 62(1): 296-304, http://dx.doi.org/10.1109/TR.2013.2241197.
8. Leong Gan C, Classe F, Hashim U. Superior performance and reliability of copper wire ball bonding in laminate substrate based ball grid array. Microelectronics International 2013; 30(3): 169-175, http://dx.doi.org/10.1108/MI-11-2012-0074.
9. Liu W, He G. Storage life of silicone rubber sealing ring used in solid rocket motor. Chinese Journal of Aeronautics 2014; 27(6): 1469-1476, http://dx.doi.org/10.1016/j.cja.2014.10.013.
10. Lu C J, Meeker W O. Using degradation measures to estimate a time-to-failure distribution. Technometrics 1993; 35(2): 161-174, http://dx.doi.org/10.2307/1269661.
11. Marioli M, Meneghini M, Rossi F, et al. Degradation mechanisms and lifetime of state-of-the-art green laser diodes. physica status solidi 2015; 212(5): 974-979, http://dx.doi.org/10.1002/pssa.201431714.
12. Meeker W Q, Escobar L A , and Lu C J. Accelerated degradation tests: modeling and analysis. Technometrics 1998; 40(2): p. 89-99, http://dx.doi.org/ 10.2307/1270643.
13. Naikan V N A, Rathore A. Accelerated temperature and voltage life tests on aluminium electrolytic capacitors: a DOE approach. International Journal of Quality & Reliability Management 2016; 33(1): 120-139, http://dx.doi.org/10.1108/IJQRM-12-2014-0201.
14. Purnell P. Interpretation of climatic temperature variations for accelerated ageing models. Journal of materials science 2004; 39(1): 113-118, http://dx.doi.org/10.1023/B:JMSC.0000007734.71945.93.
15. Si X S, Wang W, Hu C H, et al. Remaining useful life estimation based on a nonlinear diffusion degradation process. IEEE Transactions on Reliability 2012; 61(1): 50-67, http://dx.doi.org/10.1109/TR.2011.2182221.
16. Tajima K, Hotta H, Yamada Y, et al. Accelerated test on electrochromic switchable mirror based on magnesium alloy thin film in simulated environment of various relative humidities. Solar Energy Materials and Solar Cells 2012; 99: 76-83, http://dx.doi.org/10.1016/j.solmat.2011.06.018.
17. Takeda E, Suzuki N. An empirical model for device degradation due to hot-carrier injection. IEEE Electron Device Letters 1983; 4(4): 111-113, http://dx.doi.org/10.1109/EDL.1983.25667.
18. akulov N V, Myshlyavtsev A V, Malyutin V I. Estimation of in-use Guaranteed Rubber Lifetime test methods. Procedia Engineering 2015; 113: 479-483, http://dx.doi.org/10.1016/j.proeng.2015.07.339.
19. Van Noortwijk J M, Pandey M D. A stochastic deterioration process for time-dependent reliability analysis. Proceedings of the Eleventh IFIP WG 7.5 Working Conference on Reliability and Optimization of Structural Systems 2004; 259-265.
20. Wang C H, Hsu C H. New nonisothermal Arrhenius temperature integral approximate formula. Environmental Engineering Science 2012; 29(10): 964-971, http://dx.doi.org/10.1089/ees.2011.0292.
21. Wang Z, Shang S, Zhai G, et al. Research on storage degradation testing and life prediction based on ARMA and wavelet transform model for aerospace electromagnetic relay. IEEE 60th Holm Conference on Electrical Contacts (Holm) 2014; 1-8, http://dx.doi.org/10.1109/HOLM.2014.7031021.
22. Xiao K, Gu X, Peng C. Reliability evaluation of the O-type rubber sealing ring for fuse based on constant stress accelerated degradation testing. Journal of Mechanical Engineering 2014; 16: 62-69, http://dx.doi.org/10.3901/JME.2014.16.062.
23. Yang G. Accelerated life tests at higher usage rates. IEEE Transactions on Reliability 2005; 54(1): 53-57, http://dx.doi.org/10.1109/TR.2004.841730.
24. Yu H F. Designing an accelerated degradation experiment by optimizing the estimation of the percentile. Quality and Reliability Engineering International 2003; 19(3): 197-214, http://dx.doi.org/10.1002/qre.518.

25. Yu H F. Designing an accelerated degradation experiment with a reciprocal Weibull degradation rate. Journal of statistical planning and inference 2006; 136(1): 282-297, http://dx.doi.org/10.1016/j.jspi.2004.06.030.

Li SUN
Xiao-Hui GU
College of Mechanical Engineering
Nanjing University of Science and Technology
No. 200, Xiaolingwei Street, Nanjing, China

Pu SONG
Science and Technology on Combustion and Explosion Laboratory
Xi'an Modern Chemistry Research Institute
No. 168, Zhangba East Street, Xi'an, China

Yi DI
College of Mechanical Engineering
Nanjing University of Science and Technology
Nanjing 210094, China

E-mails: huixi_alice@163.com, gxiaohui@njust.edu.cn, songpu73@163.com, detiancai007@163.com

Article citation info:
STĘPIEŃ S, SZAJNAR S, JASZTAL M. Problems of military aircraft crew's safety in condition of enemy counteraction. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 441–446, http://dx.doi.org/10.17531/ein.2017.3.15.

Sławomir STĘPIEŃ
Stanisław SZAJNAR
Michał JASZTAL

PROBLEMS OF MILITARY AIRCRAFT CREW'S SAFETY IN CONDITION OF ENEMY COUNTERACTION

PROBLEMY BEZPIECZEŃSTWA ZAŁOGI WOJSKOWEGO STATKU POWIETRZNEGO W WARUNKACH PRZECIWDZIAŁANIA PRZECIWNIKA*

The presented paper consists outline of the probabilistic method of evaluation of military aircraft crew's safety, which took into consideration enemy counteraction. The specific attention was focused on estimation of durability of ejection seat, which is a means of pilot's emergency escape from aircraft. The basis of the presented model is probability of pilot's danger to life for single sortie caused by enemy. Formulated differentiation equation characterises process of increment of successful sortie number. The equation after transformation into partial differential equation served for establishing of successful sortie distribution function and subsequently for calculation of crew safety indicators.

***Keywords**: military aircraft crew, safety, ejection seat, durability, probability.*

W artykule przedstawiono zarys probabilistycznej metody oceny bezpieczeństwa załogi wojskowego statku powietrznego uwzględniającej niszczące działanie przeciwnika. Największą uwagę skupiono na szacowaniu trwałości fotela katapultowego jako środka do awaryjnego opuszczania samolotu przez pilota. Podstawą prezentowanego modelu jest prawdopodobieństwo powstania zagrożenia dla życia pilota w pojedynczym locie statku powietrznego spowodowane przeciwdziałaniem przeciwnika. Sformułowano równanie różnicowe charakteryzujące w ujęciu probabilistycznym proces przyrostu liczby udanych lotów bojowych statku powietrznego. Równanie to po przekształceniu w równanie różniczkowe cząstkowe, posłużyło do wyznaczenia funkcji rozkładu udanych lotów bojowych, a następnie wskaźników bezpieczeństwa załogi.

***Słowa kluczowe**: załoga wojskowego statku powietrznego, bezpieczeństwo, fotel katapultowy, trwałość, prawdopodobieństwo.*

1. Introduction

Military aircraft crew's safety is a very complex problem. The complexity of the problem emerges from necessity to rescue pilot's life in dynamic emergency condition, when time for making a decision is short [1, 8-9, 16-17,]. Therefore, element of "human factor", cannot be overlooked [3, 14]. World literature on this issue has an interdisciplinary nature and relates to these matters in a general way, often purely descriptive. In aviation, pilot's life rescue in emergency situation is realised with use of special devices which are called ejection seats. They are mounted only on military aircraft, what can restrict an access to information about their usage. Hence, specialised literature on this issue is not too extensive.

In Poland, studies of a practical and application nature in these field were mainly carried out in the Air Force Institute of Technology and years ago also in the Institute of Aviation, however analytical studies were performed in the Military University of Technology, Military Institute of Aviation Medicine and Warsaw University of Technology as well as other research centres worked on aerospace technology e.g. Rzeszow University of Technology.

World literature consist publication related to different areas of ejection process. Some of them raise anthropometric question connected with ejection [4], medical question as well as conditioning which may occur during ejection process [2, 3, 7, 10-14, 22-24]. Publications relating to collected combat experience also occur [13, 22]. One of the line of research is computer modelling of ejection process [5, 6, 9, 15, 25, 26], which enable to simulate various scenarios of the process, analysis of conditions of safety escape from an aircraft as well as examination of factors that affect pilot's organism in different situations of emergency escape. Huge diversity of the publications confirms interdisciplinarity of the subject area.

In the domestic literature, major part of the papers concern medical problems of ejection process and was published by scientists from the Military Institute of Aviation Medicine [23, 24]. Operating and maintenance manuals for particular type of ejection seats are also available. However, papers such as that on quantitative evaluation of crew safety with regard to ejection seat usage, are extremely rare [18].

In today's world, various types of ejection seats are employed in different types of military aircraft. The ejection seats differ in their design solution, dimensions (size, weight), established stabilisation system, survival equipment and accessory. [17, 19]. The direct inspiration for researching in the field of ejection seats development was necessity of perpetual modernisation of military aircraft towards performance and effectiveness improvement, which in turn involves indispensability of emergency escape subsystem perfection [21]. In the course of modernization, the question arises: how to assess pilot's safety for military aircraft equipped with particular rescue system?

Currently, about the idea of what type of ejection seat is installed in particular aircraft decides more often than not: type of an aircraft (who is a manufacturer), tactical and technical capabilities as well as

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

economic conditions. It also happens that, various types of ejection seat manufactured by different producers are installed on the same type of aircraft. Furthermore, each of producers have in his tender from several to over a dozen types of ejection seats, resulting significant diversity of possible technical solutions [19]. For an analysis of military aircraft crew safety, disparity of tasks in time of peace and war must also be count. Hence, further questions need answers: How to evaluate aircraft crew safety? What tool for such assessment apply? Thus, authors of present paper proceeded to study of the method (technique) which allows such an assessment, and is helpful in utilisation planning as well as aircraft and their rescue systems modernisation.

Presented method bases on probabilistic calculation, which take into consideration aircraft and ejection seat reliability. Carried out solutions do not exhaust the problem, but represents distinct progress in the quest for answers to the questions raised and provide the inspiration for formulation further questions connected with ejection seat usage for aircraft crew life rescue.

Ejection seats represents very complex technical objects, which are essential elements of the aircraft equipment and enable pilot to escape from an aircraft in very short time in emergency situation. So, it was assumed that pilot's safety during flight depends on:

1) aircraft reliability;
2) ejection seat reliability;
3) pilot health and calmness in emergency situation [1].

Additionally, in case of enemy counteraction, following elements have influence on pilot's safety:

1) military aircraft vulnerability;
2) ability to self-defense;
3) effectiveness of enemy impact on aircraft in order to destroy it during fighting;
4) quality of air traffic organisation (dynamic of operation) and crew's skills [18].

Pilot's safety was considered under the following assumptions:

1) on airfield (military base) possibility of aircraft destruction (as well as ejection seat) is negligible;
2) aircraft is destructed by enemy during combat mission;
3) probability of aircraft destruction during one sortie equals Q, and ejection seat remain operable (exist possibility to utilise the seat);
4) ejection seat durability is the same as aircraft durability.

The leading role in pilot's safety evaluation plays aircraft durability, which is defined as a number of sorties or flying time until destruction of aircraft in combat condition. „Lifetime" of the ejection seat can be measured by the number of the aircraft successful flights until destruction. The main subject of the presented paper was determination of the ejection seat durability in the above mentioned scope and assessment of crew's safety indicators.

2. Method of durability assessment with use of difference equation

This chapter presents method of determination of distribution of successful flight' number for assumed time range or for the whole aircraft's durability in terms of fighting. With the distribution of the number of successful flights considered one tried to calculate interesting parameters. The process of increment of the number of successful flight was considered as a function of time or number of sorties.

It was assumed that the sorties occurs randomly with a certain intensity λ. Therefore, for the range Δt following condition is fulfil:

$$\lambda \Delta t \le 1,$$

where: Δt can be considered as a duration time of one sortie. Above mentioned condition states that aircraft do not fly continuously but there are exist random intervals between consecutive sorties. For the sake of completeness, it was assumed that probability of aircraft's destruction during one sortie equals Q.

Let $U_{z,t}$ denote the probability that for the moment t the number of successful flights is z. For the applied symbols, the increment of the number of flights will be described in probabilistic way by the following difference equation:

$$U_{z,t+\Delta t} = (1-\lambda\Delta t)U_{z,t} + \lambda\Delta t(1-Q)U_{z-1,t}\,. \tag{1}$$

The difference equation (1) means what fallows: the probability that for the time $t+\Delta t$ the number of successful flights amount to z is equal to the sum of probabilities of the following events:

- flight did not happen during Δt and aircraft have accomplished z successful flights until the moment t,
- successful flight was accomplished during Δt, aircraft' destruction did not happened and aircraft have accomplished z-1 successful flights until the moment t.

After rearranging to the functional notation:

$$u(z,t+\Delta t) = (1-\lambda\Delta t)u(z,t) + \lambda\Delta t(1-Q)u(z-1,t), \tag{2}$$

where: $u(z,t)$ is a probability density function of the successful flights' number for the moment t.

Taylor's expansion was used for above mentioned equation.

$$u(z,t+\Delta t) = u(z,t) + \frac{\partial u(z,t)}{\partial t}\Delta t + \frac{1}{2!}\frac{\partial^2 u(z,t)}{\partial t^2}\Delta t^2 + \frac{1}{3!}\frac{\partial^3 u(z,t)}{\partial t^3}\Delta t^3 + \frac{1}{4!}\frac{\partial^4 u(z,t)}{\partial t^4}\Delta t^4 + \ldots$$

$$u(z-1,t) = u(z,t) - \frac{\partial u(z,t)}{\partial z} + \frac{1}{2!}\frac{\partial^2 u(z,t)}{\partial z^2} - \frac{1}{3!}\frac{\partial^3 u(z,t)}{\partial z^3} + \frac{1}{4!}\frac{\partial^4 u(z,t)}{\partial z^4} + \ldots$$

For t it is two elements of expansion, and for z three elements. After transformation and rearrangement of the equation (2), there was obtained:

$$\frac{\partial u(z,t)}{\partial t} = -\lambda Q u(z,t) - \lambda(1-Q)\frac{\partial u(z,t)}{\partial z} + \frac{1}{2}\lambda(1-Q)\frac{\partial^2 u(z,t)}{\partial z^2}\,. \tag{3}$$

Let introduce the notations: $c = \lambda Q : b = \lambda(1-Q) : a = \lambda(1-Q)$.

Coefficient c indicates aircraft' destruction intensity. Then coefficient b and a despite the same description have somewhat different meaning, b is an average increment of the number of successful flights in the time range Δt, instead a is an average square of increment of the number of successful flights in the time range Δt. Convergence of the descriptions for a and b result from the fact that in the time range Δt can occur increase only by one flight, so an increment and square of increment are equal. Then:

$$\frac{\partial u(z,t)}{\partial t} = -c u(z,t) - b\frac{\partial u(z,t)}{\partial z} + \frac{1}{2}a\frac{\partial^2 u(z,t)}{\partial z^2}\,. \tag{4}$$

To present solution of equation (4) authors made use of equation's solution of the Fokker-Planck type [21] in the following form:

$$\frac{\partial \bar{u}(z,t)}{\partial t} = -b\frac{\partial \bar{u}(z,t)}{\partial z} + \frac{1}{2}a\frac{\partial^2 \bar{u}(z,t)}{\partial z^2}. \tag{5}$$

We search for the solution of the particular equation (5), which, by $t \to 0$ is concurrent to the so-called Dirac function: $\bar{u}(z,t) \to 0$ for $z \neq 0$ and $\bar{u}(0,t) \to \infty$, but in this way that the integral of function $\bar{u}(z,t)$ equals a unity for all $t > 0$. For above mentioned condition, the solution of equation (5) takes the form [21]:

$$\bar{u}(z,t) = \frac{1}{\sqrt{2\pi at}} e^{-\frac{(z-bt)^2}{2at}}. \tag{6}$$

Having taken account of these considerations one can present particular solution of equation (4) which takes the form:

$$u(z,t) = ce^{-ct}\bar{u}(z,t). \tag{7}$$

In order to verify the correctness of the solution, following transformations have been made:

$$\frac{\partial u(z,t)}{\partial t} = -c^2 e^{-ct}\bar{u}(z,t) + ce^{-ct}\frac{\partial \bar{u}(z,t)}{\partial t} = \left[-c - \frac{1}{2t} + b\frac{(z-bt)}{at} + \frac{(z-bt)^2}{2at^2}\right]u(z,t),$$

$$\frac{\partial u(z,t)}{\partial z} = ce^{-ct}\frac{\partial \bar{u}(z,t)}{\partial z} = -\frac{(z-bt)}{at}u(z,t),$$

$$\frac{\partial^2 u(z,t)}{\partial z^2} = ce^{-ct}\frac{\partial^2 \bar{u}(z,t)}{\partial z^2} = \left[-\frac{1}{at} + \frac{(z-bt)^2}{a^2t^2}\right]u(z,t).$$

Having put above relationships into (4), one gets:

$$\left[-c - \frac{1}{2t} + b\frac{(z-bt)}{at} + \frac{(z-bt)^2}{2at^2}\right]u(z,t) = \left[-c - b\left[-\frac{(z-bt)}{at}\right] + \frac{1}{2}a\left[-\frac{1}{at} + \frac{(z-bt)^2}{a^2t^2}\right]\right]u(z,t).$$

Then, after arrangement:

$$\left[-c - \frac{1}{2t} + b\frac{(z-bt)}{at} + \frac{(z-bt)^2}{2at^2}\right]u(z,t) = \left[-c - \frac{1}{2t} + b\frac{(z-bt)}{at} + \frac{1}{2}\frac{(z-bt)^2}{at^2}\right]u(z,t).$$

As seen, left side of the relationships is equal to right side, which proves the correctness of this solution.

Finally, the distribution of the number of aircraft successful flights was obtain:

$$u(z,t) = ce^{-ct}\frac{1}{\sqrt{2\pi at}} e^{-\frac{(z-bt)^2}{2at}}. \tag{8}$$

Function (8) has features of density function since:

$$\int_0^\infty \int_{-\infty}^\infty u(z,t)\,dz\,dt = 1.$$

Having determined the distribution of the number of aircraft successful flights (density function (8)) it is possible to obtain:

1) average value of the number of aircraft successful flights:
   a) for the lifetime;
   b) for the finite period of time;
2) for established number of successful flights z_1:
   a) probability that the number of successful flights is less than or equal to z_1;
   b) probability that the number of successful flights is greater than z_1 as a function of time.

Expected value of the number of successful flights for time less than t_1:

$$E[z] = \int_0^{t_1}\int_{-\infty}^\infty zu(z,t)\,dz\,dt = \frac{b}{c}\left[1-(1+ct_1)e^{-ct_1}\right] = \frac{(1-Q)}{Q}\left[1-(1+\lambda Q t_1)e^{-\lambda Q t_1}\right]. \tag{9}$$

If we take into account longer time period, i.e. time $t_1 \to \infty$, we received well known relationship which describe average value of the number of successful flights for the aircraft' lifetime:

$$E_T[z] = \frac{(1-Q)}{Q}. \tag{10}$$

Hence, the number of successful flights for finite range of time is described by equation (9).

Probability, that the number of successful flights is less than or equal to z_1 for time t_1 with possibility of aircraft' destruction is represented by:

$$P_{z_1}^{(1)}(t_1) = \int_0^{t_1}\int_{-\infty}^{z_1} u(z,t)\,dz\,dt. \tag{11}$$

Probability, that the number of successful flights is greater than z_1 for time t_1 with possibility of aircraft' destruction is represented by:

$$P_{z_1}^{(2)}(t_1) = \int_0^{t_1}\int_{z_1}^{\infty} u(z,t)\,dz\,dt. \tag{12}$$

Probability, that in the range of time (0, t_1) aircraft will not be destroyed has following form:

$$\begin{aligned} P^{(3)}(t_1) &= 1 - \left(P_{z_1}^{(1)}(t_1) + P_{z_1}^{(2)}(t_1)\right) = 1 - \int_0^{t_1} ce^{-ct}dt = \\ &= 1 - \left[-e^{-ct}\right]_0^{t_1} = 1 - \left[-e^{-ct_1} + 1\right] = e^{-ct_1} \end{aligned} \tag{13}$$

It can be demonstrated that specified probabilities bring the total number to one.

$$P_{z_1}^{(1)}(t_1) + P_{z_1}^{(2)}(t_1) + P^{(3)}(t_1) = 1.$$

The equation (13) can be written down in the form:

$$P^{(3)}(k) = e^{-Qk}, \tag{14}$$

where: $k = \lambda t_1$ - number of sorties performed until time t_1.

Probability that in the range of number of sorties *(0, k)* aircraft will be destroyed adopts the following form:

$$Q_k(k) = 1 - e^{-Qk}. \tag{15}$$

3. Outline of the pilot's safety assessment

In case of rise to loss of the ship hazard, pilot for the sake of saving one's own life is forced to trigger ejection seat. The success of ejection process depends mainly on the following factors [18]:

1) time to reach a decision about ejection;
2) course of ejection (including pilot's landing after ejection process);
3) conditions of the ejection process;
4) type of aircraft and type of ejection seat;
5) behaviours and skills of pilot during ejection process.

In real situations time for reaching decision about emergency escape is predominantly very short. Additionally, in this kind of situation pilot often tries to remedy the threat to aircraft. Great influence on making the right decision about ejection has "human factor" and other factors which determine pilot's mental state.

Taking decision about ejection necessitates the implementation of a series of activities which influence ejection process. These activities are more or less automated. Reliable performance of the operations has significant influence on the results of ejection. As mentioned in the introduction above, reliability of the emergency escape depends on aircraft's and ejection seat's type. Emergency escape does not always end successfully, to a large extent the results depends on conditions in which it occurred. Probabilistic evaluation of the pilot's safety can be, depending on the assumptions and simplifications, more or less accurate.

In this paper, model is limited by taking into account aircraft and ejection seat reliability. Reliability of an aircraft and ejection seat, by virtue of assumptions, are considered as a separate sets of events:

$$R_S(t) + Q_S(t) = 1, \tag{16}$$

$$R_F(\tau) + Q_F(\tau) = 1, \tag{17}$$

where: $R_S(t)$ - aircraft reliability in the time range *(0, t)*;

$Q_S(t)$ - unreliability i.e. probability of the aircraft destruction for the time range *(0, t)*;

$R_F(\tau)$ - ejection seat reliability at the time of its use;

$Q_F(\tau)$ - ejection seat unreliability at the time of its use.

The above equation (16) refers to the aircraft, and equation (17) to the ejection seat. Using formulas (16) and (17), probability of pilot's survive in the time range *(0, t)* can be determined in the following form:

$$\bar{R} = R_S(t) + Q_S(t) R_F(\tau), \tag{18}$$

where: $\bar{R}$ - probability that, in the time range *(0, t)*, aircraft was not destroyed or the destruction of an aircraft occurred and pilot survived thanks to use of the operable ejection seat.

Probability of loss of pilot's life during sorties in the time range *(0, t)* takes the form:

$$\bar{Q} = Q_S(t) Q_F(\tau), \tag{19}$$

where: $\bar{Q}$ - probability of loss of pilot's life during sorties in the time range *(0, t)*.

Dependency (18) shows that pilot's safety depends on reliability of an aircraft and reliability of ejection seat. Dependencies (18) and (19), how easy it is to check, bring the total number to one, which proved the correctness of the above formulas for the assumptions made. Equation (18) for exponential distribution as a function of number of sorties has following form:

$$\bar{R}_k = e^{-Qk} + \left(1 - e^{-Qk}\right) R_F(\tau), \tag{20}$$

where: k – number of sorties.

4. Illustration of the calculations

Due to lack of available and reliable data relating to combat use of aircraft, which are indispensable for evaluation of required results, below hypothetical data were used.

Input data for the calculations:

– sorties intensity $\lambda = 2$ [$^1/_{day}$];
– probability of aircraft destruction during one sortie $Q = 0{,}05$ [-].

The input data illustrate situation, where sorties are curried out twice daily, and probability of aircraft destruction during one sortie equals 0,05.

Figure 1 presents changes over time of an average value of the number of successful flights $E[z]$ (calculated according to (9)) and dashed line presents stationary value $E_T[z]$ pursued by $E[z]$ (calculated according to (10)). $E_T[z] = 19$.

While, figure 2 presents changes over time of probabilities $P_{z_1}^{(1)}(t_1)$, $P_{z_1}^{(2)}(t_1)$, $P^{(3)}(t_1)$ calculated according to the equations (11), (12) and (13).

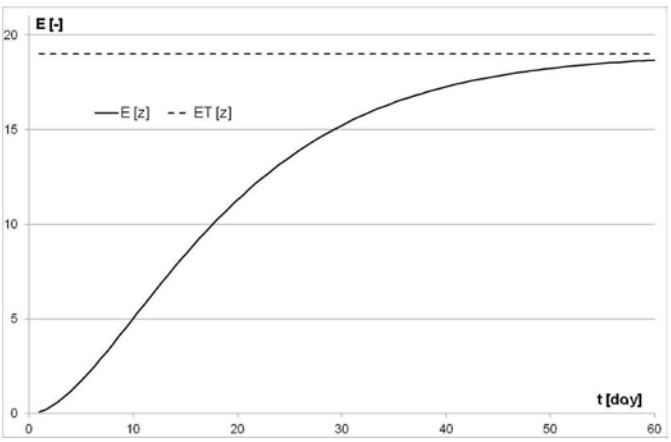


Fig. 1. Average value of the number of successful flights

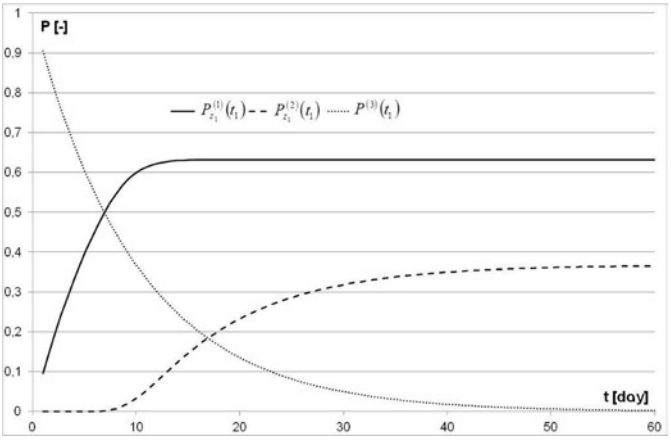


Fig. 2. Graphs of calculated probabilities: $P_{z_1}^{(1)}(t_1)$ – probability, that the number of successful flights is less than or equal to z_1=19 for time t_1, $P_{z_1}^{(2)}(t_1)$ - probability, that the number of successful flights is greater than z_1=19 for time t_1, $P^{(3)}(t_1)$ - probability, that in the range of time (0, t_1) aircraft will not be destroyed

5. Summary

The requirements of the modern battlefield in the field of aircraft engineering forces us to seek reliable (not only intuitive) answers to important questions: what effect should be expected on battlefield as a result of activity of specified type of one's own or enemy aircraft, and how aircraft engineering should be formed in order to achieve a goals for assumed probability and under particular conditions. Furthermore, very important aspect of effectiveness evaluation of military aircraft usage is assessment of aircraft durability and crew safety during combat mission [20]. It seems that presented method of crew safety evaluation in terms of the enemy counteracting can be used for the preliminary assessment of pilot safety with specific rescue system applied to the aircraft. Additionally, the method supports decision-making during combat mission as well as facilitates obtaining required indicators in the field of safety and reliability for combat use of aircraft.

With the use of distribution of the number of successful flights obtained in presented work it is possible to determine average value of the number of successful flights (as presented at figure 1) as well as probability of achievement established number of successful flights and probability of aircraft endurance for specified time (figure 2). Presented numerical example shows possible utilitarian aspects of use the method outlined in the work.

References

1. Borgoń J. Niezawodność i bezpieczeństwo systemu pilot - statek powietrzny. Informator Instytutu Technicznego Wojsk Lotniczych 1987; 269/87.
2. Chiou W Y, Ho B L, Kellogg D L. Hazard potential of ejection with canopy fragmentation. Aviation, Space, and Environmental Medicine 1993; 64(1): 9-13.
3. Davis J R, Johnson R, Stepanek J, Fogarty J A (Editors). Fundamentals of Aerospace Medicine. 4rd Edition. Lippincott, Philadelphia: Williams & Wilkins, 2008.
4. Edwards M. Anthropometric measurements and ejection injuries. Aviation, Space, and Environmental Medicine 1996; 67(12): 1144-1147.
5. Głowiński S, Krzyżyński T. Modelling of the ejection process in a symmetrical flight. Journal of Theoretical and Applied Mechanics 2013; 51(3): 775-785.
6. Grzesik N, Czapla R. Aircraft crew escape system assistant. Safety and Reliability: Methodology and Applications. CRC Press 2014: 791-796.
7. Hearon B F, Thomas, H A, Raddin J H. Mechanism of vertebral fracture in the F/FB-111 ejection experience. Aviation, Space, and Environmental Medicine 1982; 53(5): 440-448.
8. Lewis M E. Survivability and injuries from use of rocket-assisted ejection seats: analysis of 232 cases. Aviation, Space, and Environmental Medicine 2006; 77(9): 936-943.
9. Maryniak J, Maryniak A, Ładyżyńska-Kozdraś E, Folte U. Katapultowanie - możliwości, problemy i modelowanie. Nauka, Innowacje, Technika 2004; 5 (7): 28-45.
10. McBratney C M, Rush S, Kharod C U. Pilot ejection, parachute, and helicopter crash injuries. Journal of special operations medicine 2014; 14(4): 92-94.
11. Moiseev I B, Strakhov A I, Churilov I K, Vovkodav V S, Radchenko S N. Medical outcomes of emergency ejections from Russian aircrafts in 2003-2010. Aviation, Space, and Environmental Medicine 2014; 48(2): 57-62.
12. Newman D G. Survival outcomes in low-level ejections from high performance aircraft. Aviation, Space, and Environmental Medicine 2013; 84(10):1061-1065, https://doi.org/10.3357/ASEM.3626.2013.
13. Osborne RG, Cook AA. Vertebral fracture after aircraft ejection during Operation Desert Storm. Aviation, Space, and Environmental Medicine 1997; 68(4): 337-341.
14. Rainford D J, Gradwell D P (Editors). Ernsting's Aviation and Space Medicine. 5th Edition. CRC Press 2006.
15. Ramm A G, Kaleps I. Modeling of the ejection process. Mathematical and Computer Modelling 1994; 20: 95-101, https://doi.org/10.1016/0895-7177(94)90221-6.
16. Szajnar S W. Diagnozowanie w podsystemie opuszczania statku powietrznego. Informator Instytutu Technicznego Wojsk Lotniczych 1995; 330/95: 291-298.
17. Szajnar S W. Ocena bezpieczeństwa i modelowanie w systemach awaryjnego opuszczania samolotu wojskowego. Warszawa: Wojskowa Akademia Techniczna, 2014.

18. Szajnar S W, Tomaszek H. Problemy wyznaczania wskaźników trwałości fotela katapultowego i bezpieczeństwa załogi wojskowego statku powietrznego w warunkach działań bojowych. Zagadnienia Eksploatacji Maszyn 2001; 3(127): 95-104.
19. Szajnar S W, Wojtkowiak M. Problemy bezpieczeństwa załogi statku powietrznego w sytuacjach awaryjnych". Warszawa: BIL-GRAF, 1999.
20. Tomaszek H, Wróblewski M. Podstawy oceny efektywności eksploatacji systemów uzbrojenia lotniczego. Warszawa: Dom Wydawniczy Bellona", 2001.
21. Tomaszek H, Żurek J, Jasztal M. Prognozowanie uszkodzeń zagrażających bezpieczeństwu lotów statku powietrznego. Radom: Wydawnictwo Naukowe Instytutu Technologii Eksploatacji, 2008.
22. Williams C S. F-16 pilot experience with combat ejections during the Persian Gulf War. Aviation, Space, and Environmental Medicine 1993; 64(9): 845-847.
23. Wojtkowiak M. Adaptacja ustroju do działania przyspieszeń w katapultowaniu rzeczywistym i pozorowanym. Lekarz Wojskowy 1971; 1:40-45.
24. Wojtkowiak M. Wpływ pozycji pilota na urazy kręgosłupa podczas katapultowania. Medycyna Lotnicza1973; 42: 5-15.
25. Yu J, Lin G, WU M. Numerical Simulation of Deceleration Performance of Ejection Seat. Acta Aeronautica et Astronautica Sinica 2006; 27(6): 1033-1038.
26. Yu J; Lin G, Mao X. Numerical Simulation of Ejection Seat and Analysis of Performance Under Adverse Attitudes. Acta Aeronautica et Astronautica Sinica 2010; 31(10): 1927-1932.

Sławomir STĘPIEŃ
Stanisław SZAJNAR
Michał JASZTAL
Military Academy of Technology
ul. Gen. Sylwestra Kaliskiego 2, 00–908 Warsaw, Poland

E-mails: sstępień@wat.edu.pl, sszajnar@wat.edu.pl, mjasztal@wat.edu.pl

Article citation info:
HE L, LI G, XING L, CHEN Y. An autonomous multi-sensor satellite system based on multi-agent blackboard model. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 447–458, http://dx.doi.org/10.17531/ein.2017.3.16.

Lei HE
Guoliang LI
Lining XING
Yingwu CHEN

AN AUTONOMOUS MULTI-SENSOR SATELLITE SYSTEM BASED ON MULTI-AGENT BLACKBOARD MODEL

AUTONOMICZNY WIELOCZUJNIKOWY SYSTEM SATELITARNY OPARTY NA WIELOAGENTOWYM MODELU TABLICOWYM

Traditional Earth observation satellite cannot work well in terms of emergencies, environmental uncertainties and scientific events discovery. Therefore, it is of significance to study the new generation of autonomous Earth observation satellite. In order to develop an autonomous satellite system with distributed and coordinated functions, this paper proposes an autonomous satellite system based on distributed multi-agent blackboard model. Multiple agents including functions of pre-processing, planning, scheduling and execution are designed. Agents share information and communicate through a blackboard which stores the task sequence, the action sequence and the satellite status. An adaptive rule-based heuristic scheduling algorithm and a forward search planning algorithm are proposed. The simulation experiments and computational results prove that the system can deal with scientific events discovery, satellite faults, cloud obscuration and emergencies without human intervention, which can greatly enhance the efficiency and reliability of Earth observation satellites. The validity of the proposed model and algorithm is proved.

***Keywords**: autonomous satellite, blackboard model, multi-agent, adaptive scheduling, planning.*

Tradycyjne satelity obserwacji Ziemi nie nadają się do pracy w sytuacjach kryzysowych, warunkach niepewności środowiskowej oraz w okolicznościach związanych z odkryciami naukowymi. Dlatego też istotne znaczenie ma badanie nowej generacji autonomicznych satelitów obserwacji Ziemi. W celu opracowania autonomicznego systemu satelitarnego o rozproszonych i skoordynowanych funkcjach, w niniejszej pracy zastosowano rozproszony wieloagentowy model tablicowy. Zaprojektowano agenty, w tym funkcje wstępnego przetwarzania, planowania, harmonogramowania i wykonania. Agenty te wymieniają między sobą informacje i komunikują się za pośrednictwem tablicy (ang. blackboard), na której przechowywane są informacje dotyczące sekwencji zadań i działań oraz stanu satelity. Zaproponowano adaptacyjny, regułowy, heurystyczny algorytm harmonogramowania oraz algorytm planowania metodą wyszukiwania w przód. Przeprowadzone eksperymenty symulacyjne oraz wyniki obliczeń dowodzą, że omawiany system sprawdza się w przypadkach odkryć naukowych, awarii satelitarnych, zachmurzenia oraz w sytuacjach kryzysowych nie wymagając interwencji człowieka, co może znacznie zwiększać wydajność i niezawodność satelitów obserwacji Ziemi. W pracy wykazano trafność proponowanego modelu i algorytmów.

***Słowa kluczowe**: satelita autonomiczny, model tablicowy, wieloagentowy, harmonogramowanie adaptacyjne, planowanie*

1. Introduction

The Earth observation satellite (EOS) is becoming one of the major platforms for spatial information access, whose task is to get remote sensing data of specified targets on the surface of Earth according to the users' observation requests. EOS has the advantages of wide range and long imaging time, and is not restricted by national boundaries. Therefore, EOS is applied in many fields including disaster monitoring [21] and information support [15]. EOS is regarded as an important part of a country's comprehensive national power. The development of EOS has attracted the attention from all over the world.

The problem of EOS task planning and scheduling is mainly to study the resource allocation of multiple satellites and sensors, scheduling and planning multiple tasks, producing the action sequences and uploading to the satellites, i.e., deciding what the satellite should do at a certain time. Its working flow chart is shown in Fig. 1.

Currently, almost all the command and control instructions for satellite systems are issued by the ground control station and most of satellite tasks are time driven. The data acquired by the satellites

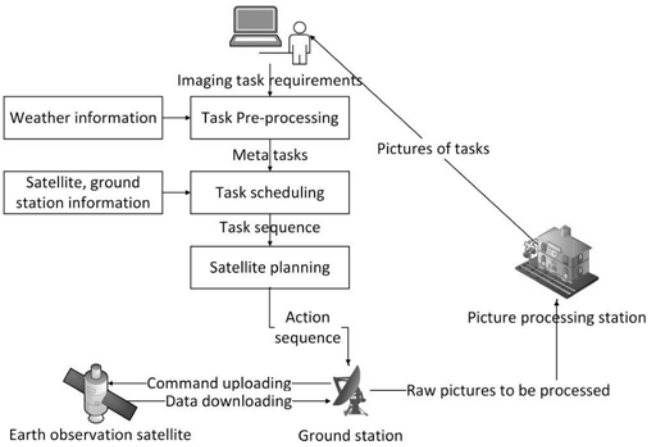


Fig. 1. Earth observation satellite management mode

is also downloaded, processed and analysed by the ground stations. With the increasing number of the demands for the space information, this management mode is facing a series of challenges [17]:

- The satellite system needs the ability to deal with emergencies within a short period of time. However, due to the presence of satellite-ground communication delay and the limited communication bandwidth, the response time is too long to respond the emergencies and other uncertain changes;
- The deployment of ground stations is limited by geographical conditions and national boundaries;
- The number of satellites and the amount of information to be downloaded both increase dramatically, which brings difficulties to the control, communication, coordination and ground operations of the whole system.

Therefore, a new satellite management mode is necessary. An effective mode is to reduce the human intervention and make the satellite system more intelligent, so that it can work independently to complete the tasks. With the development of satellite technology, a new generation of satellite has a certain capability of computing and processing, which makes the development of autonomous satellite become the focus of research in the satellite field [10]. Compared with the traditional satellite, the autonomous satellite includes the following advantages: higher performance, better fault tolerance, higher reliability and lower maintenance cost. The system can deal with the uncertainties caused by environmental changes. The response time for emergencies will be shorten. Science events, such as floods and volcano eruptions, can be found by the satellite. The system also has high reconfigurability and extendibility.

There are several autonomous satellite in service currently. The US National Aeronautics and Space Administration (NASA) EO-1 [6] is an autonomous satellite and is able to discover science events on the Earth surface. Because of this autonomous management, EO-1 can save more than one million dollars every year and the events discovered annually worth more than 1.8 million dollars [7]. EO-1 uses SVM [4] to classify land, ice, cloud, etc. The software system on EO-1 is called ASE [24], and uses a scheduling method called CASPER. CASPER is quite fast and uses a simple iterative repairing algorithm to search for feasible solutions [24]. US Air Force's TacSat-3 [22] is another autonomous satellite designed to allow end users to send requests to the satellite directly. The satellite can respond to these requests automatically. FireBIRD is an autonomous satellite developed by German Aerospace Center (DLR) to discover and monitor high-temperature events such as forest fires [21]. However, the planning and scheduling processes are mostly done on the ground. The on-board part only optimizes the results and gives feedbacks according to the on-board information [18]. So, the application of FireBIRD is quite limited. French PLEIADES project started in 2000 and two autonomous satellites (Pleiades 1A and 1B) were launched in 2011 and 2012 respectively [12]. Pleiades has a powerful on-board autonomous planning ability. The planning process is regarded as a decision problem and solved by the constraint network on timelines (CNT) structure [3].

There are also some theoretical research. In 2010, European Space Agency (ESA) developed a distributed multi-agent system in its Distributed Agents For Autonomy (DAFA) study [5] and showed the advantages of performing on-board autonomy using agent technology in the satellite system. Araguz et al. [1] proposed a distributed satellite system to decompose the complex planning problem into several local planners. And they are coordinated by a global planner. The resource exchange management is not considered and there are no experimental results. Golkar's Federated Satellite Systems (FSS) [13] has similar ideas. They both use distributed modules to deal with complex missions but the FSS is mainly designed for constellations with multiple satellites.

From the above analysis, it can be seen that the current autonomous satellite systems are mostly mission-specific and the extendibility is low. And distributed modules and decomposition of tasks are two main trends of autonomous satellite development. Considering the advantages and disadvantages of these systems, this paper has designed a single autonomous multi-sensor satellite system based on the distributed blackboard model, which is a multi-agent system (MAS), to provide a reference for the research of the future autonomous satellite.

The remainder of this paper is organized as follows: Section 2 introduces the structure of this autonomous satellite system based on the blackboard model. Section 3 describes the mathematical model of the agile satellite observation task scheduling. Section 4 proposes an adaptive rule-based heuristic scheduling algorithm and a forward search planning algorithm. In section 5, simulation experiments are presented. Section 6 concludes the paper.

2. The system structure based on the blackboard model

2.1. Blackboard model

The first blackboard model was proposed for a speech understanding system [8]. The blackboard model is used to coordinate the relationship of each agent in the system, controlling the establishment of agents, communication and cooperation.

The basic idea of the blackboard model is as follows. A complex problem will be solved by multiple experts together, and the blackboard is a shared working space. All experts can see the blackboard. When the problem and initial data are written on the blackboard, the solving process starts. All the experts monitor the blackboard and when one finds the information on the blackboard is enough to support his work, he will start calculation and then write the results on the blackboard. The new information can be used by other experts. The process goes on until a final result is obtained and the complex problem is solved.

The blackboard model is widely used for establishing multi-agent intelligent systems [2, 20, 23], especially for the ones with dynamic, uncertain and complex tasks. It provides a flexible and efficient method of communication among agents. The agents have different solving skills and work independently. Because the different modules do not interact with each other, the blackboard model is suitable for the system that already has some existing software. Adding new functions to this structure is quite convenient. And for this satellite system, some functions (such as task execution) already exist in a traditional satellite. Therefore, the blackboard model is quite suitable for this autonomous satellite system.

2.2. System structure design

As shown in Fig. 2, the structure realizes a distributed management of the functions of the satellites. The sub-systems have independent functions but are connected with each other to complete the imaging tasks.

The blackboard model is a BDI model, including three basic sets: the belief set, the desire set and the intention set. The three sets refer to the current resources the agents can use, the tasks the agents want to complete and the actions must be performed to complete the tasks respectively.

The blackboard module stores the status information of all the agents, including the task queue, the resource queue, belief and task execution evaluation results, which are the initial solution, partial solution, final solution and other system status. The function plug-ins "read from/write on" the blackboard by the subscriber and the releaser. All plug-ins exchange information through the blackboard and adjust the information on the blackboard. The task queue includes tasks

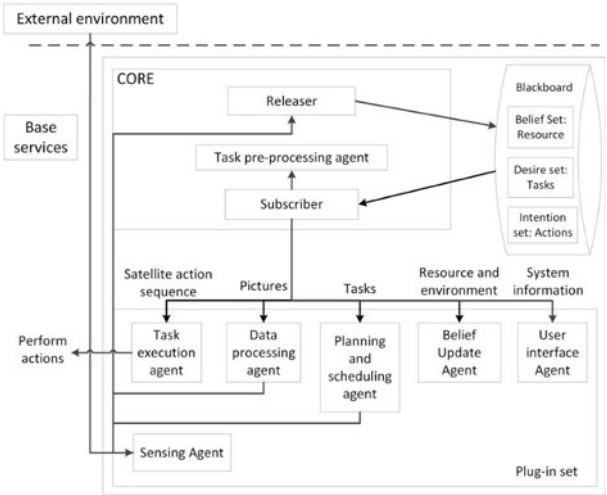


Fig. 2. Multi-agent system structure

to be planned which form the desire set *D* of agents. The resource queue records the status of all the resources of the satellite. The status of these resources forms the belief set *B* of the agents. The planning and scheduling agent is used to generate the final task execution plan, which is the action sequence of the satellite. And this sequence forms the intention set *I* of the agents. The satellite completes tasks through the implementation of the actions.

The core module is used to provide the necessary basic functions for the satellite. Every single autonomous satellite has the same core structure.

The agent plug-in set includes some basic plug-ins which are the agents. These agents are also called knowledge source (KS). These agents are the function modules of the autonomous satellites, determining all the functions that can be used to complete a certain task. New agents of plug-ins can be added to extend the abilities of the satellites to meet the different demands. All the agents solve the problem related to their functions and work independently.

Subscriber and releaser: The subscriber is a monitoring mechanism of the blackboard model. It can select and activate a proper KS (agent) according to the information on the blackboard, to make the KS respond to the changes of information. Releaser is used to write the partial results from KSs on the blackboard.

The external environment: An autonomous satellite should have the ability to sense the environment and use the information to adjust plans and find new events on-line. We assume that this multi-sensor satellite carries a cloud detector, a visible optical sensor and an infrared sensor for observation, and there are some other sensors for monitoring the status of the satellite. The external environment has the following features:

- Partly-accessible: The cloud detector on the satellite can partly sense the cloud information in a few minutes, such as the location of clouds;
- Non-deterministic: Due to the real-time changes of environment and the complexity of the outer space, the implementation of an action can bring about different results, which also shows the importance of the autonomy;
- Non-episodic: When the satellite is executing a task, it should consider not only the current episode, but also the continuously changing environment as well as the information from other agents, and sometimes some simple predictions are also needed. Therefore, it is non-episodic;
- Dynamic: The environment is highly variable, such as clouds moving and various dynamic changes in the outer space;
- Continuous: The change of the environment is continuous and not discrete.

Therefore, the external environment of the satellite is very complex. Many uncertainties need to be considered.

There are many agents supporting each kind of functions in this satellite system. The function of each agent is as follows:

(1) Pre-processing agent: This agent can calculate the visible time windows (VTW) of each task according to the satellite track forecast and the locations of the targets. Here, VTW means the time period when the target is visible for the satellite. It is also known as the meta-task. Within the VTWs, the satellite can observe the targets and complete the imaging process.
(2) Planning and scheduling agent: This agent schedules the tasks according to the belief set and desire set and produces a task sequence with timestamps. According to the sequence, the satellite action sequence is then produced. This process is called planning. At the same time, this agent can also deal with some abnormal conditions to improve the reliability and reduce the maintenance cost.
(3) Belief agent: It is used to sense the status of resources and the satellite faults.
(4) Execution agent: Sensors and other devices on the satellite are controlled by this agent. At the same time, the latest information in the execution process is generated. For example, if a task is not executed as planned, such as delaying several seconds, the following task execution plan should be adjusted accordingly.
(5) Data processing agent: After an image is acquired, the agent can analyse it and check whether the target is successfully imaged. If not, it can generate a new task to image the target again.
(6) Sensing agent: This agent is used to sense the external environment and discover new tasks. By sensing the external environment, the impact of environmental uncertainties on imaging can be detected. The satellite can perform actions accordingly. The reliability can be increased. And new tasks of science events can be discovered.
(7) User interface agent: This agent is used to monitor the blackboard information. When the content of the blackboard changes, the interface is updated to facilitate the users to monitor the system. This agent can process the orders from the ground. Besides, the agent also allows us to upgrade and reconstruct the different modules on the satellite, such as adding a more efficient scheduling algorithm. This mechanic can extend the abilities of the satellite.

In a cycle of this system, the agents work as follows:

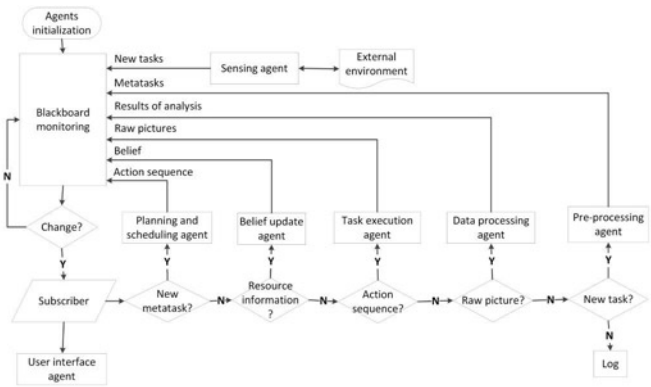
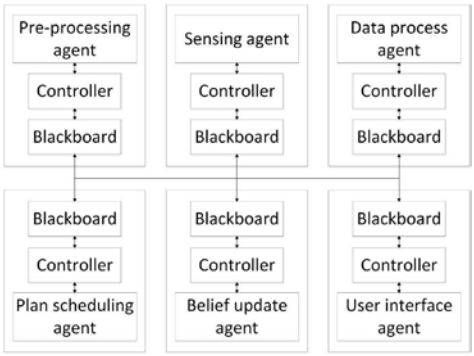
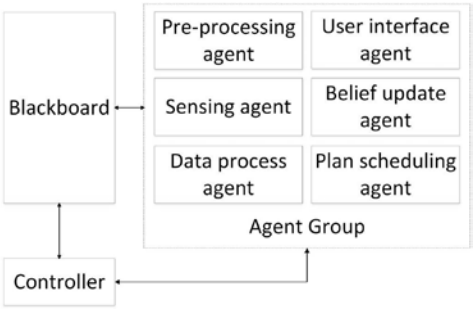


Fig. 3. Workflow of this system

Fig. 3 shows the relationship between the agents and the blackboard. The subscriber acts as a role of monitoring mechanism of the content on the blackboard. When the corresponding data is available, it dynamically selects and activates the appropriate agent.



(a) Distributed structure



(b) Centralized structure

Fig. 4. Comparison of two kinds of hardware structures

Fig. 4 shows the hardware structure of this satellite system. In this paper, the autonomous satellite system adopts a distributed structure, which is shown in Fig. 4(a). To show the characteristics of this structure, we also give the centralized structure as a comparison in Fig. 4(b). There is only one blackboard in the system in the logic and software level (shown in Fig. 2), `while in the physical` and hardware `level`, there is one local blackboard in each agent, forming a distributed blackboard structure. This structure enables parallel and distributed computing of multiple agents on different units to cooperate and complete tasks. The local blackboards keep information consistency by the network broadcast. The contents in all the local blackboards are identical. Therefore, the agents in this system have a shared blackboard in the logic level, which provides data and intermediate results for the agents. Because the agents are distributed in different processors, all agents can work independently. The failure of one agent does not affect the ability of other agents, which can improve the reliability of the system. By contrast, in the centralized structure, there is only one blackboard in the physical level. All the agents work depending on the only one blackboard, making it difficult to realize the parallel computing of different agents. Besides, if the blackboard fails, the agents cannot work anymore. The reliability of this centralized structure is lower.

2.3. Pseudo code

The pseudo code of the planning and scheduling agent is shown to demonstrate the working principle.

Pseudo code 1: Planning and scheduling agent

```
public class SchedulePlan extends Agent {
  Releaser rel = new Releaser();
  Subscriber sub = new Subscriber();
  protected void setup() {
    this.addBehaviour(new CyclicBehaviour() {
      public void action() {
        ACLMessage msg = receive();
        if (msg != null) {
          tasklist = sub.readtask();
          timelist = sub.readVTW();
          tasklist = schedule(tasklist,timelist);
          actoutlist = plan(tasklist);
          rel.writeschedule(tasklist);
          rel.writeplan(actoutlist);
          ACLMessage msg1 = new ACLMessage(ACLMessage.INFORM);
          msg1.setContent("action sequence ready");
          send(msg1);
        }
      }
    });
  }
```

3. Mathematical model of agile satellite task scheduling problem

The main parameters involved in the mathematical model of agile satellite task scheduling are listed as follows:

- $T = \{t_1, t_2, \cdots, t_{N_T}\}$: task set, where N_T represents the total number of tasks;
- P_i : $P_i \in [1,10]$, the priority of task t_i , whose maximum is 10 and minimum is 1;
- N_i : the number of VTWs of task t_i ;
- t_{ik}^0 , t_{ik}^{mid} , t_{ik}^{end} : the start time, middle time and end time of the kth VTW of task t_i ;
- g_i : the reward of task t_i ;
- d_i : the length of the observation window, i.e. the time required for imaging the task t_i ;
- t_{ij}^{tran}, t_{ij}^{stable}: the transition time and stable time needed from task t_i to task t_j .
- Decision variable:
- x_{ik}: $x_{ik} \in \{0,1\}$, if the kth VTW of task t_i is chosen to observe the task, $x_{ik} = 1$; otherwise, $x_{ik} = 0$;
- to_i^0, to_i^{end} : the start time and end time of the observation window.

The objective function is to maximize the total reward of all the tasks:

$$\max P_T = \sum_{i=1}^{N_T} \sum_{k=1}^{N_i} x_{ik} g_i \tag{1}$$

s.t.

$$\forall t_i \in T, \sum_{k=1}^{N_i} x_{ik} \le 1 \tag{2}$$

$$x_{ik} = 1, \ t_{ik}^0 \le to_i^0 < to_i^{end} \le t_{ik}^{end} \tag{3}$$

$$to_i^{end} + t_{ij}^{tran} + t_{ij}^{stable} \le to_j^0 \tag{4}$$

$$to_i^{end} - to_i^0 = d_i \tag{5}$$

Equation (2) means that each task can only be observed at most once. Equation (3) is the VTW constraint, restricting that the imaging time of each task must be inside the VTW of the task. Equation (4) is the transition time constraint, meaning that the imaging of the previous task does not affect the beginning of the latter task. This is the time-dependent characteristic of agile satellite scheduling. The transition time depends on the attitude of the satellite, which depends on the observation time of two successive observations. Equation (5) is the observation time constraint. The imaging must last for some time in order to ensure the integrity of the picture.

According to the characteristics of the agile satellite imaging process, the image with the best imaging quality can be acquired when the satellite is right above the target, i.e., when it is acquired at the middle of the VTW. Therefore, the reward of the task t_i can be calculated by the following equation:

$$g_i = \frac{P_i}{10}\left[10 - \frac{9\left|t - t_{ik}^{mid}\right|}{t_{ik}^{mid} - t_{ik}^{0}}\right] \tag{6}$$

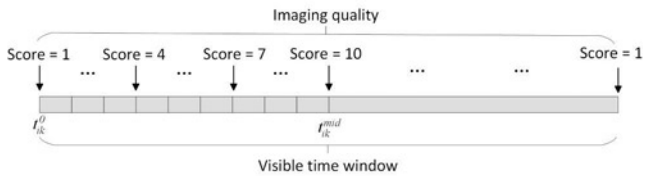


Fig. 5. The influence of imaging time on the imaging quality

As shown in Fig. 5, the highest score for the imaging quality is 10. And according to equation (6), 10 can be acquired in the middle of the VTW. When the picture is imaged at the two margins, the score is 1. Finally, the reward of a task can be calculated considering the priority and the imaging quality, as shown in equation (6).

4. Autonomous satellite system algorithm design

4.1. Adaptive Rule-based heuristic scheduling algorithm

The scheduling problem of agile satellite is a typical NP-Hard combinatorial optimization problem [16], whose solution space is very large, and with the increasing number of the tasks, the complexity of the problem grows exponentially. Besides, because the computing resources on the satellite are limited, the scheduling and planning processes cannot be too complex, otherwise it will influence the execution of the plan. Therefore, this paper uses an adaptive rule-based heuristic algorithm (ARHA) to generate a feasible and near optimal schedule.

We have designed four operators to sort the tasks to be scheduled. They are:

- Priority sorting: tasks with higher priority will be considered preferentially.
- Opportunity sorting: tasks with fewer VTWs will be chosen first because those with more VTWs are easier to schedule.
- Observation time sorting: this operator chooses tasks whose VTWs are earlier.
- Conflict sorting: the operator sort the tasks by a descending order of the overlapping degree. Here the overlapping degree means the length of time one VTW overlaps with other VTWs. Because the tasks with higher overlapping degree may have conflict with other tasks, they might reduce the total reward.

To increase the adaptability of this algorithm, we have designed an adaptive mechanism. Each operator has a score and a weight and both of them depend on the performance of the operators. We regard every 50 times of scheduling as a section. At the end of each section, the weight for each operator is updated. The score is updated in every time of scheduling and initialized at the end of each section. Every time an operator is chosen, its score is update according the following equation:

$$s_i' = s_i + \gamma \tag{7}$$

where, s_i and s_i' are the current score and updated score respectively, and γ is the reward rate, which equals to the proportion of the total reward divided by the total priority. At the end of each section, the weights of operators are updated:

$$\omega_i' = (1-\lambda)\omega_i + \lambda s_i / \sum_{i=1}^{4} s_i \tag{8}$$

where, ω_i is the historical weight and $\lambda \in [0,1]$ is a reaction factor controlling the sensitiveness of the weights.

Roulette wheel mechanism is used to choose the operators. And the probability one operator can be chosen is calculated as follows:

$$p_i = \omega_i / \sum_{j=1}^{4} \omega_j \tag{9}$$

The flow of the algorithm is as follows:

Step 1: Get the task information and the satellite VTW data;

Step 2: Choose a sorting operator using the roulette wheel mechanism. Generate a task sequence according to this operator. If one operator cannot determine the order, choose another operator until the tasks are sorted;

Step 3: Set the counter $k = 0$;

Step 4: Find all VTWs of task t_k;

Step 5: Find the n VTWs with the smallest overlapping degree;

Step 6: Calculate the best imaging quality of each VTW;

Step 7: Choose the VTW with the highest imaging quality;

Step 8: Update the scores of all the operators. If this is the end of a section, update the weights of all the operators and initialize the scores;

Step 9: If $k = \mathrm{n}$, or all the tasks are successfully scheduled, or the maximum iteration time is met, the algorithm terminates; otherwise, go to Step 3.

The flow chart of scheduling algorithm is shown in Fig. 6:

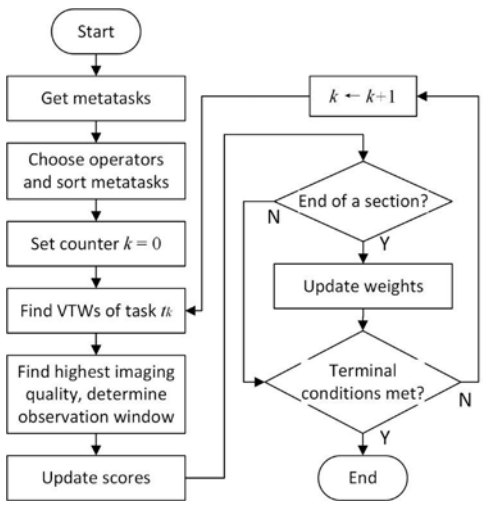
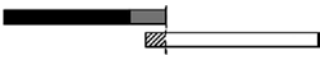
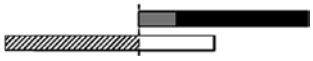


Fig. 6. Flow chart of scheduling algorithm

4.2. Cutting the VTWs

Before determining the best imaging quality of a VTW, the VTW should be cut in order to meet the transition time constraint. By judging the relationship between the determined observation time window and the VTW, the VTW can be cut accordingly. Assuming that task t_j has been determined for observation. And k^{th} VTW of task t_i can be cut according to the observation time window of task t_j. As shown in Table 1, the VTW is cut according to the six different relationships.

Table 1. Cut the VTWs Remaining part Occupied window Maneuver time Cut window

ID	Relationship	Diagram
1	$to_j^0 < t_{ik}^0 < t_{ik}^{end} < to_j^{end}$	
2	$t_{ik}^0 < to_j^0 < to_j^{end} < t_{ik}^{end}$	
3	$to_j^{end} < t_{ik}^0$	
4	$t_{ik}^{end} < to_j^0$	
5	$to_j^0 < t_{ik}^0 < to_j^{end} < t_{ik}^{end}$	
6	$t_{ik}^0 < to_j^0 < t_{ik}^{end} < to_j^{end}$	

4.3. A forward search algorithm based on STRIPS

The satellite action sequence can be generated according to the task sequence produced in the above section. A forward search algorithm based on STRIPS [9] is used. The STRIPS operator can describe an action as a process from a previous state to a subsequent state. Typically a STRIPS operator consists of the following three parts:

- Set PC: the preconditions of the operator. An action can be performed only when all the preconditions in PC are met;
- Set D: delete list. It shows the states changed when an action is performed;
- Set A: add list. It includes the changes brought by an action.

According to the above description, a typical satellite action, such as turning on the camera, can be described by STRIPS as the following format:

Pseudo code 2: Action "turn on the camera" STRIPS

```
TurnOnCamera()
PC: CameraOff()∧CameraPreheated()∧Ontime()
D: CameraOff()
A: CameraOn()∧PictureCaptured(task.duration)
```

Therefore, for the action of turning on the camera, the preconditions include: the camera is off, the camera preheating has been finished and the imaging time is due; delete list includes the state that the camera is off; the add list includes: the camera is on and the photo is imaged. The preconditions of the next action "turn off the camera" are PictureCaptured () and Ontime (). Thus, when the image is captured, the camera will be turned off.

As shown above, if an action's preconditions are all met, this action will be performed. Using a forward search algorithm, the action sequence can be generated.

5. Simulation experiments and results discussion

A laptop with Intel Core i5-3317U CPU 1.70GHz, 8G memory and Windows 10 system is used to develop and run the system. The developing platform is JDK 1.8+ NetBeans IDE 8.1 + Jade 4.0.

5.1. Online events simulation

In order to show the system's autonomy characteristics, this paper has designed four kinds of online events:

5.1.1. The discovery of new tasks

If there is no task on the blackboard, the system will be in a waiting state. By clicking the user interface button "new task", we simulate the process that the sensing agent (Sense) discovers a new task and the task is written on the blackboard.

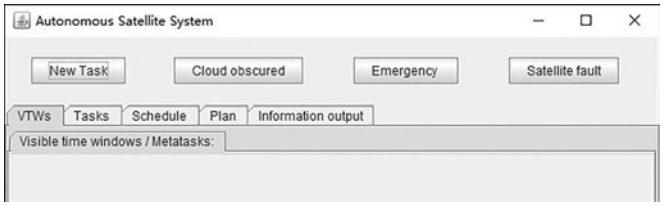
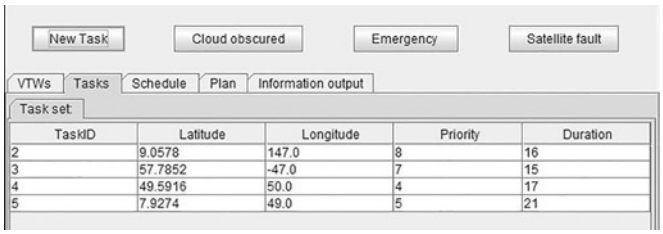
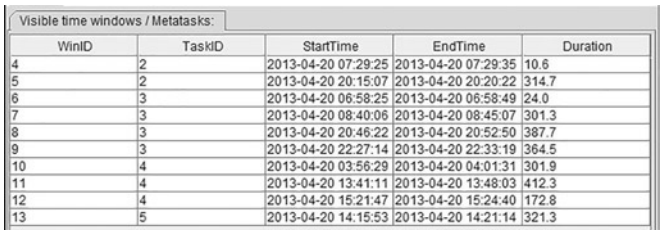


Fig. 7. Autonomous satellite system interface



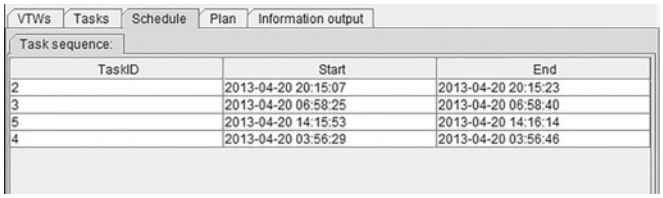
TaskID	Latitude	Longitude	Priority	Duration
2	9.0578	147.0	8	16
3	57.7852	-47.0	7	15
4	49.5916	50.0	4	17
5	7.9274	49.0	5	21

(a) Task list



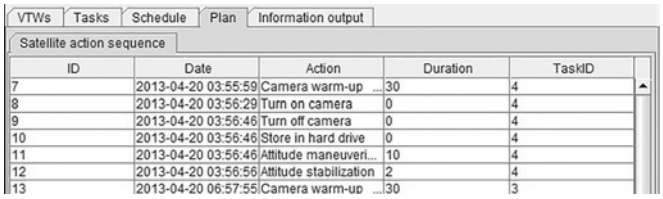
WinID	TaskID	StartTime	EndTime	Duration
4	2	2013-04-20 07:29:25	2013-04-20 07:29:35	10.6
5	2	2013-04-20 20:15:07	2013-04-20 20:20:22	314.7
6	3	2013-04-20 06:58:25	2013-04-20 06:58:49	24.0
7	3	2013-04-20 08:40:06	2013-04-20 08:45:07	301.3
8	3	2013-04-20 20:46:22	2013-04-20 20:52:50	387.7
9	3	2013-04-20 22:27:14	2013-04-20 22:33:19	364.5
10	4	2013-04-20 03:56:29	2013-04-20 04:01:31	301.9
11	4	2013-04-20 13:41:11	2013-04-20 13:48:03	412.3
12	4	2013-04-20 15:21:47	2013-04-20 15:24:40	172.8
13	5	2013-04-20 14:15:53	2013-04-20 14:21:14	321.3

(b) VTW list



TaskID	Start	End
2	2013-04-20 20:15:07	2013-04-20 20:15:23
3	2013-04-20 06:58:25	2013-04-20 06:58:40
5	2013-04-20 14:15:53	2013-04-20 14:16:14
4	2013-04-20 03:56:29	2013-04-20 03:56:46

(c) Schedule / task sequence with timestamp

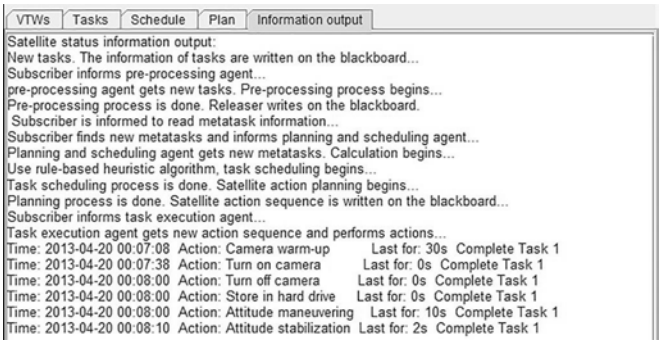


ID	Date	Action	Duration	TaskID
7	2013-04-20 03:55:59	Camera warm-up ...	30	4
8	2013-04-20 03:56:29	Turn on camera	0	4
9	2013-04-20 03:56:46	Turn off camera	0	4
10	2013-04-20 03:56:46	Store in hard drive	0	4
11	2013-04-20 03:56:46	Attitude maneuveri...	10	4
12	2013-04-20 03:56:56	Attitude stabilization	2	4
13	2013-04-20 06:57:55	Camera warm-up ...	30	3

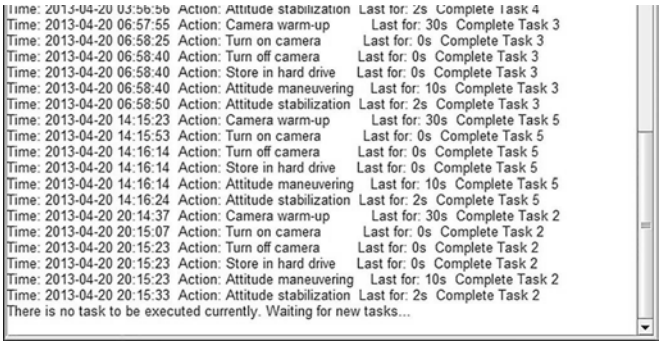
(d) Plan / satellite action sequence

Fig. 8. System screenshots for the new task discovery

Fig. 7 shows the interface of this system. The above four buttons, respectively, represent the four types of events. The VTW window is used to display a list of all the VTWs whose tasks are not executed.



(e) System information after task 1 is executed



(f) System information after all tasks are executed

Fig. 8. (continued) System screenshots for the new task discovery

When one task is executed, all the corresponding VTWs will be deleted from the blackboard. The task window lists all the tasks that are not executed; schedule window and plan window show the schedule list and action list that are not performed respectively. The information window is used to display the information of the satellite.

Here we assume that the satellite finds five new tasks. Fig. 8(a)-(e) show the screenshots after the first task is executed.

Fig. 8(a) shows that task 1 is removed from the blackboard after the execution. Fig. 8(b)-(d) show that all the VTWs, task sequence and action list about task 1 are removed from the blackboard. From Fig. 8(e) we can monitor the satellite actions and all the information when the satellite is processing the event. Fig. 8(f) is the screenshot after all the five tasks are executed.

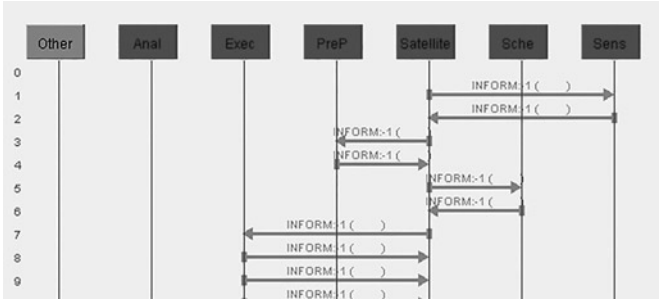


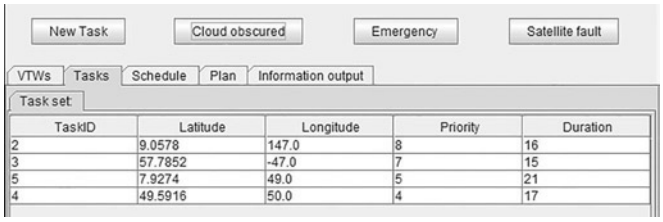
Fig. 9. Information flow between Agents

Fig. 9 shows the information flow between agents. "Anal" is the data processing agent. "Exec" is the execution agent. "PreP" is the pre-processing agent. "Satellite" is the user interface agent. "Sche" is the planning and scheduling agent. "Sens" is the sensing agent. After the discovery of new tasks, the agents cooperate with each other. The whole process of pre-processing, scheduling, planning, execution and picture analysis are performed automatically without human intervention.

5.1.2. The image is obscured by clouds

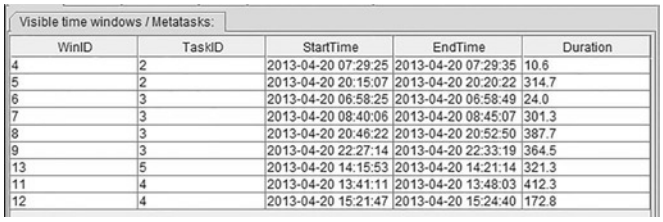
In the real situation, after imaging a target, the picture will be analyzed by the data processing agent to check the usability of the picture. In our simulation, after imaging a task, the ID of this task will be recorded. If the system detects "cloud obscured" button is clicked, the last picture imaged is obscured by the clouds.

Here, after task 4 is executed, we click the "cloud obscured" button. Fig. 10(a)-(c) are the screenshots:



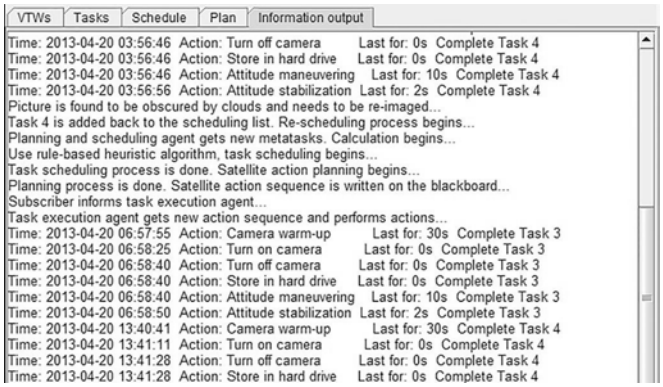
TaskID	Latitude	Longitude	Priority	Duration
2	9.0578	147.0	8	16
3	57.7852	-47.0	7	15
5	7.9274	49.0	5	21
4	49.5916	50.0	4	17

(a) Task list

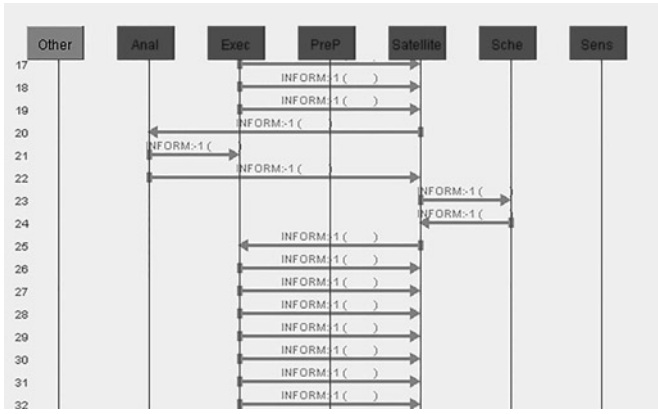


WinID	TaskID	StartTime	EndTime	Duration
4	2	2013-04-20 07:29:25	2013-04-20 07:29:35	10.6
5	2	2013-04-20 20:15:07	2013-04-20 20:20:22	314.7
6	3	2013-04-20 06:58:25	2013-04-20 06:58:49	24.0
7	3	2013-04-20 08:40:06	2013-04-20 08:45:07	301.3
8	3	2013-04-20 20:46:22	2013-04-20 20:52:50	387.7
9	3	2013-04-20 22:27:14	2013-04-20 22:33:19	364.5
13	5	2013-04-20 14:15:53	2013-04-20 14:21:14	321.3
11	4	2013-04-20 13:41:11	2013-04-20 13:48:03	412.3
12	4	2013-04-20 15:21:47	2013-04-20 15:24:40	172.8

(b) VTW list



(c) System information after task 4 is obscured



(d) Information flow for cloud obscuration

Fig. 10. Screenshots for the cloud obscuration

From Fig. 10(a), it can be seen that task 4, which is already executed, is added back to the blackboard. In Fig. 10(b), the VTWs of task 4 is added back too. But the VTW around 3:56 is deleted because the time has passed. In Fig. 10(c), we can find that after task 4 was ex-

ecuted at around 3:56, the data processing agent find that task 4 is obscured by clouds. Therefore, the picture is deleted from the hard drive and task 4 should be re-imaged, triggering the on-board re-scheduling process. Then task 4 is re-imaged at around 13:41 according to the results of the re-scheduling process.

Fig. 10(d) shows the information flow between agents. We can see that the data processing agent (Anal) sends two message, informing task execution agent to wait and informing user interface agent to update the user interface. Because task 4 is added back to the blackboard, the planning and scheduling agent are activated.

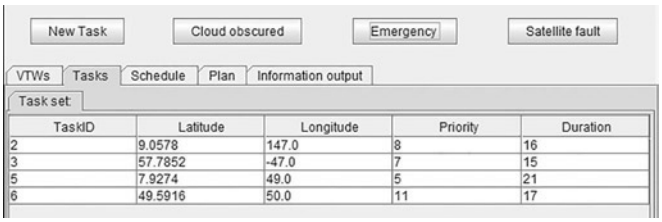
To sum up, after the picture is obscured by the cloud, the agents trigger the online re-scheduling process, delete the useless picture and re-image task 4. The storage is saved and useful data can be downloaded, increasing the reward rate and reliability.

5.1.3. Emergency task

This button represents the situation that a task with a higher priority has arrived or been found and if this task conflicts with some determined tasks, a re-scheduling and re-planning process will be triggered and the tasks with lower priorities will be re-scheduled to meet the requirement of the task with the higher priority.

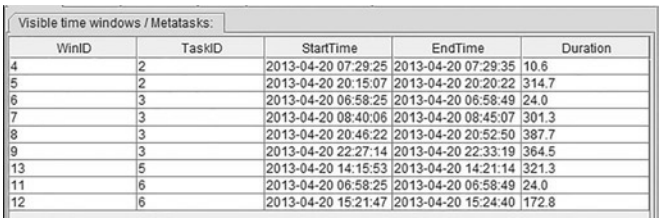
According to the original plan, after the execution of task 4, task 3 should be executed (as shown in Fig. 8(d)). Here, we add an emergency task with priority 11 after the execution of task 4 and make the new task conflict with task 3. The screenshots of the system are shown in Fig. 11(a)-(c).

In Fig. 11(a), there is a new task 6 listed in the task queue, and its priority is 11. Fig. 11(b) shows that the VTWs of task 6 conflict with those of task 3. Therefore, task 6 with higher priority should be imaged first. In Fig. 11(c), we can find that after execution of task 4, a new task with a higher priority arrives. And its VTWs conflict with those of task 3. In the re-scheduling process, the time for executing task 3 is used to execute task 6. And task 3 is postponed to be executed at 8:39.



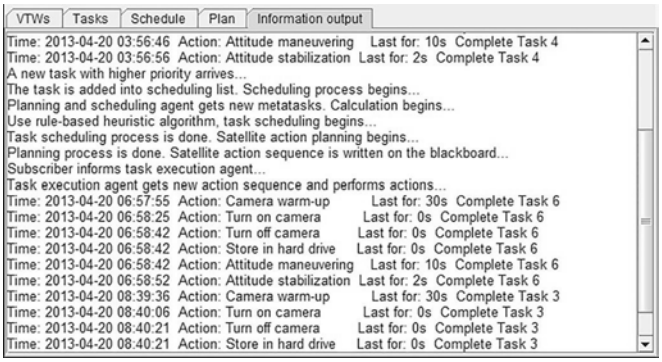
TaskID	Latitude	Longitude	Priority	Duration
2	9.0578	147.0	8	16
3	57.7852	-47.0	7	15
5	7.9274	49.0	5	21
6	49.5916	50.0	11	17

(a) Task list

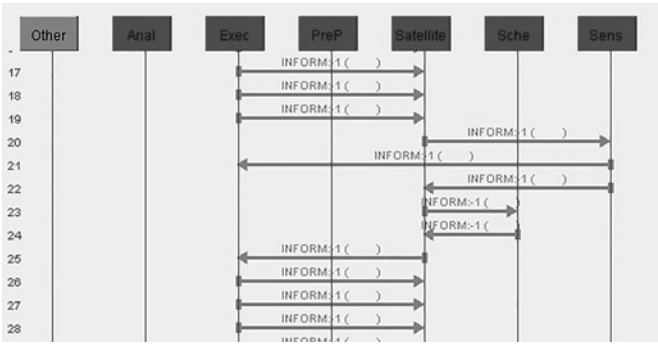


WinID	TaskID	StartTime	EndTime	Duration
4	2	2013-04-20 07:29:25	2013-04-20 07:29:35	10.6
5	2	2013-04-20 20:15:07	2013-04-20 20:20:22	314.7
6	3	2013-04-20 06:58:25	2013-04-20 06:58:49	24.0
7	3	2013-04-20 08:40:06	2013-04-20 08:45:07	301.3
8	3	2013-04-20 20:46:22	2013-04-20 20:52:50	387.7
9	3	2013-04-20 22:27:14	2013-04-20 22:33:19	364.5
13	5	2013-04-20 14:15:53	2013-04-20 14:21:14	321.3
11	6	2013-04-20 06:58:25	2013-04-20 06:58:49	24.0
12	6	2013-04-20 15:21:47	2013-04-20 15:24:40	172.8

(b) VTW list



(c) System information when emergency arrives



(d) Information flow for processing emergency

Fig. 11. Screenshots for the processing emergency

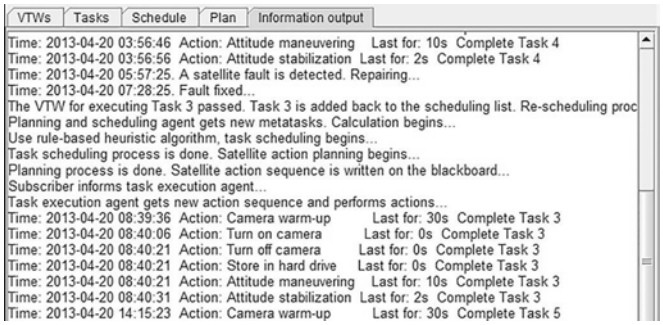
According to the original plan, after the execution of task 4, task 3 should be executed (as shown in Fig. 8(d)). Here, we add an emergency task with priority 11 after the execution of task 4 and make the new task conflict with task 3. The screenshots of the system are shown in Fig. 11(a)-(c).

In Fig. 11(a), there is a new task 6 listed in the task queue, and its priority is 11. Fig. 11(b) shows that the VTWs of task 6 conflict with those of task 3. Therefore, task 6 with higher priority should be imaged first. In Fig. 11(c), we can find that after execution of task 4, a new task with a higher priority arrives. And its VTWs conflict with those of task 3. In the re-scheduling process, the time for executing task 3 is used to execute task 6. And task 3 is postponed to be executed at 8:39.

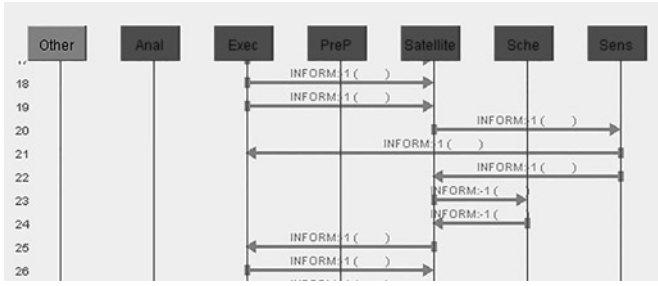
5.1.4. Satellite fault

This button simulates the existence of a satellite fault. If this button is clicked, the system will wait for five seconds (the speed of the simulation is 1000 times of the reality. Therefore five seconds in simulation is equivalent to 1.5 hours in reality), no observation task can be executed during this time period.

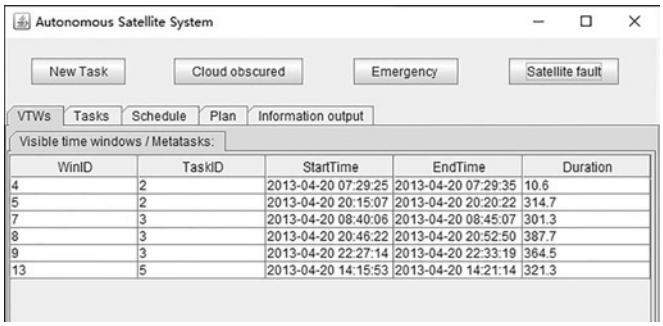
Fig. 12(a)-(b) are the screenshots after a satellite fault is detected. After the execution of task 4, a satellite fault occurs and lasts about 1.5 hours. After the fault is fixed, the VTW for imaging task 3 has passed.



(a) Task list



(b) System information when emergency arrives



WinID	TaskID	StartTime	EndTime	Duration
4	2	2013-04-20 07:29:25	2013-04-20 07:29:35	10.6
5	2	2013-04-20 20:15:07	2013-04-20 20:20:22	314.7
7	3	2013-04-20 08:40:06	2013-04-20 08:45:07	301.3
8	3	2013-04-20 20:46:22	2013-04-20 20:52:50	387.7
9	3	2013-04-20 22:27:14	2013-04-20 22:33:19	364.5
13	5	2013-04-20 14:15:53	2013-04-20 14:21:14	321.3

(c) Information flow for processing satellite fault

Fig. 12. Screenshots for processing the satellite fault

Table 2. Results of comparison of algorithms' performance

Scenario ID	Number of Tasks	Number of VTWs	Reward rate(%)			CPU time (s)		
			ARHA	CPLEX	ACO	ARHA	CPLEX	ACO
1	1	2	100	100	100	0.01	0.57	0.09
2	3	7	100	100	100	0.02	1.19	0.43
3	5	13	100	100	100	0.02	4.42	1.29
4	10	23	100	100	100	0.02	936.53	1.88
5	11	26	100	100	100	0.02	10385.10	2.42
6	12	26	100	-	100	0.03	-	2.90
7	50	117	100	-	100	0.04	-	27.57
8	100	236	100	-	98.99	80.63	-	143.66
9	150	367	100	-	100	90.62	-	307.70
10	200	466	99.49	-	98.48	612.69	-	729.68
11	250	596	100	-	100	132.51	-	1059.17
12	300	701	99.66	-	98.32	962.93	-	916.74

Therefore, a re-scheduling process is triggered to re-image task 3. In Fig. 12(b), we can find that after the fault is fixed, the re-scheduling process postpones task 3 to its second VTW (8:39) to be executed. In Fig. 12(c), the sensing agent (Sense) finds the fault and informs the task execution agent to wait and the user interface agent to update. After the fault is fixed, the planning and scheduling agent starts the re-scheduling process. The whole fault discovering and fixing process and scheduling process are conducted by the multi-agent structure, showing a high reliability of the satellite system.

5.2. Algorithm performance comparison

To test the performance of the proposed adaptive rule-based heuristic scheduling algorithm (ARHA), we compare it with the ILOG-CPLEX solver 12.6 and the ant colony optimization (ACO). ILOG-CPLEX solver 12.6 is a powerful software for solving constraint programming problem. The algorithm of the CPLEX solver is an exact algorithm, which can always find the optimal solution. ACO is an evolutionary algorithm for path optimization. We regard the sequence of tasks as the ants' path and use ACO to optimize it. We use the improved ACO proposed in paper [14]. For ARHA, the weight update parameter $\lambda = 0.5$ and the maximum iteration is 5000. For ACO, the parameters are set according to paper [14].

In this section, 12 scenarios are designed to test the performances of the above algorithms. In each scenario, a certain number of targets are randomly distributed worldwide. Each algorithm is run for five times and the average results are reported in Table 2.

For the exact algorithm of CPLEX, the solution space of this agile satellite scheduling problem is too big. We can see that the CPU running time increases exponentially with the number of tasks. And for the scenario with 12 tasks, the CPU running time becomes too long to find the optimal solution. Therefore, in practice, CPLEX cannot be used in this kind of scheduling problem, especially for this autonomous satellite online scheduling problem.

For the other two algorithms, ACO and ARHA both performs very well in terms of the reward rate. In most cases, the reward rate is 100% and ARHA is slightly better than ACO for Scenario 8, 10 and 12. As for the running CPU time, ARHA performs much better than ACO. In almost all the scenarios, ARHA uses less time to find a near optimal solution. The proposed heuristics in ARHA are helpful to guide the algorithm to find a near optimal solution in a large solution space efficiently.

The efficiency and effectiveness of ARHA for this online autonomous satellite task scheduling problem is proved.

5.3. Reliability comparison

For a traditional non-autonomous satellite, the above four kinds of events cannot be finished automatically. The satellite cannot discover events by itself. Typically, it is the ground station which receives requirements from users and then uploads the commands. And because there is no computational ability on-board and the commands uploaded are fixed, the satellite cannot deal with any uncertainties including cloud coverage, emergencies and satellite faults.

Here, we firstly compare the performance of traditional satellite and autonomous satellite under the consideration of the cloud information.

To represent the cloud information, a definition called cloud-obscuring time window (CTW) [19] is used. It refers to the time period that the target is obscured by clouds. All the ATWs should be cut according to CTWs to remove the time period that targets are obscured. Table 3 shows some examples of CTWs:

Table 3. Examples of CTWs

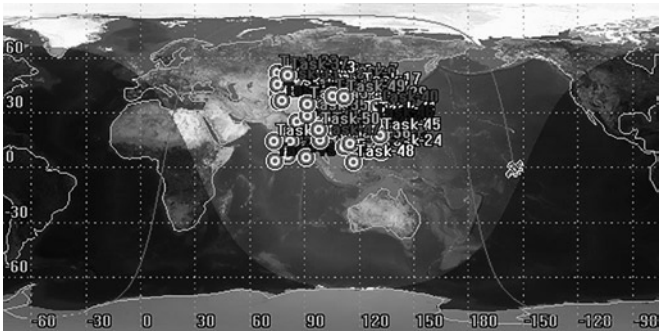
CTW ID	ATW	Start	End	Task
1	1	1:32:54	1:33:15	1
2	2	13:36:48	13:36:59	1
3	3	3:10:36	3:10:56	2
4	4	15:12:46	15:13:43	2

In Table 3, the first row shows that this CTW (1:32:54-1:33:15) obscures the ATW 1 of task 1. Then the corresponding time of ATW 1 should be removed. After all the ATWs are pre-processed, the tasks can be calculated by an off-line algorithm. For the two kinds of satellites, they have the same off-line scheduling result. And the scheduling for traditional satellite is fixed, but the autonomous satellite can re-schedule the task sequence according to the changes of the cloud information. We change the start and end time of CTWs by 2-10 seconds to simulate the movement of clouds.

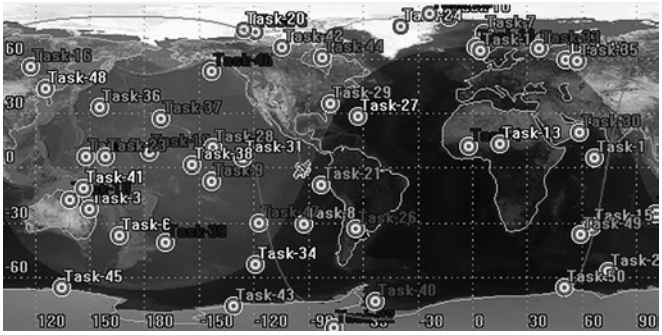
Five scenarios with two kinds of distribution modes are designed. As shown in Fig. 13(a), the targets are densely-distributed randomly in the area of China (3°N-53°N, 73°E-133°E). And Fig. 13(b) shows the distribution that targets are sparsely-distributed randomly worldwide. Table 4 listed the information of these five scenarios.

In Table 5, we use the reward rate to value the reliability of the system. It is clear that the reliability of the traditional satellite is much lower than that of the autonomous satellite. Although the traditional satellite has a high expected reward, the real reward is quite low because many pictures of the traditional satellite are obscured by clouds and the downloaded pictures are useless. By contrast, using the cloud detector, the autonomous satellite collects the real-time cloud information, re-schedules those targets that are obscured and downloads the cloud-free pictures. Therefore, the reliability of the autonomous satellite are much higher than the traditional non-autonomous satellite.

In order to show the effectiveness of the proposed distributed structure of autonomous satellite, we also compare the performance of this structure with the centralized structure of the satellite system. Compared with the centralized structure, this distributed structure has two advantages. The first one is the parallel computing of all the agents, making the online computing process fast. The other advantage is the higher



(a) Example of a densely-distributed scenario



(b) Example of a sparsely-distributed scenario

Fig. 13. Diagram of two kinds of distribution modes

Table 4. Scenario list

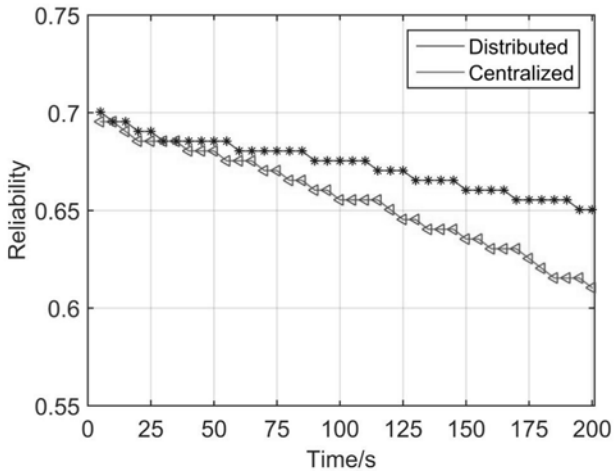
Scenario ID	Number of tasks	Tasks with ATWs	Total priority	Distribution mode
1	40	38	243	Dense
2	100	99	573	Dense
3	100	99	487	Sparse
4	150	149	863	Sparse
5	200	200	1127	Sparse

Table 5. Result list

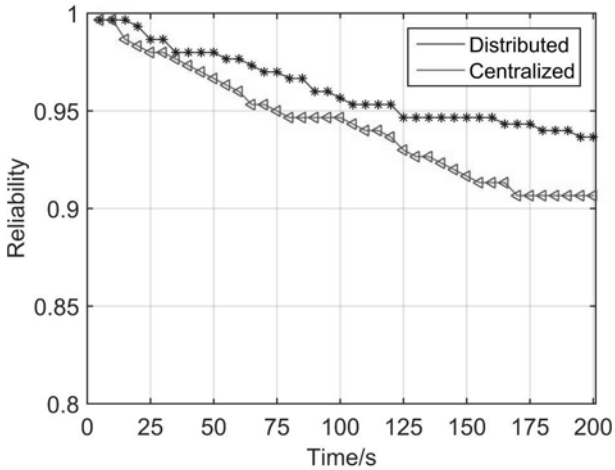
Scenario ID	Traditional satellite			Autonomous satellite	
	Expected reward	Real reward	Reliability	Reward	Reliability
1	191.22	40.61	16.71%	121.27	49.91%
2	386.11	241.40	42.13%	357.74	62.43%
3	434.44	175.02	35.94%	417.89	85.81%
4	755.95	306.90	35.56%	715.68	82.93%
5	988.04	594.78	52.78%	976.72	86.67%

reliability. Because the agents are distributed on different processing units, the failure of one unit will not affect the others.

In this experiment, we simulate the circumstance of dynamic emergence of emergency tasks. There are two scenarios. In the first scenario, there are 200 tasks that are densely distributed in the task queue and the initial reliability is 70.54%. In every five seconds, the system will add five tasks into the task queue. In the second scenario, there are 300 tasks that are sparsely distributed and the initial reliability is 100%. In every five seconds, the system will add five tasks into the task queue. The results of the reliability change with the emergence of new tasks are reported respectively in Fig. 14(a) and (b). As is shown, the distributed structure has a higher reliability for the online computing of new emergency tasks. With the emergence of



(a) Scenario 1



(b) Scenario 2

Fig. 14. The reliability change when adding new tasks to the system.

new tasks, the reliability of the centralized structure declines much faster than that of the distributed structure due to the lack of parallel computing.

In Table 6, we analyze the reliability change when different agents fail for both the distributed structure and the centralized structure. In this experiment, we design a scenario with 40 densely distributed tasks and five emergency tasks. The emergency tasks are not visible for the ground station and should be found by the satellite itself. And using this scenario, we test the reliability of the system when each of the agents fails. The result is reported in Fig. 15.

Fig. 15 The reliability of the system when each agent fails

From Fig. 15, we can see that for the centralized structure, if one of the agents fails or the blackboard fails, the satellite will become a non-autonomous satellite and the system reliability will be seriously affected. By contrast, for the proposed distributed structure, if one of the agents fails, the reliability is affected little. If the pre-processing agent and the sensing agent fail, the satellite cannot find new emergency tasks. If the plan-scheduling agent fails, the satellite cannot perform an online re-planning process so the reliability will be affected more. Due to the restricted space environment, the satellite can hardly be maintained and repaired after it is launched into space. Therefore, for the important processing unit, like the planning and scheduling agent in this paper, some redundancy can be used. The significance of other software and hardware units can also be calcu-

Table 6. The reliability analysis when different agents fail

Agent type	Distributed structure	Centralized structure
Pre-processing	The online calculation of VTWs fails. But for the tasks generated on the ground, the calculation can be done on the ground. This reliability is not affected.	All the agents are integrated on the one processing unit. If this unit fails, the whole system cannot work like an autonomous satellite. The satellite becomes a non-autonomous satellite.
Plan-scheduling	The online planning and scheduling ability fails. The satellite can still find new tasks, report and observe targets according to ground orders. The reliability is much higher than a non-autonomous satellite.	
Belief	This failure will cause a wrong estimation of satellite resource such as the battery volume, making some of scheduled tasks fail. The reliability will be affected a little.	
Data processing	Because parts of failed tasks cannot be found and deleted, some memory capacity will be wasted. The planning and scheduling are not affected.	
Sensing	The satellite cannot find new tasks by itself, but the current reliability is not affected.	
User interface	The system's reliability is not affected.	
Blackboard	-	The whole system cannot work anymore. The satellite becomes a non-autonomous satellite.

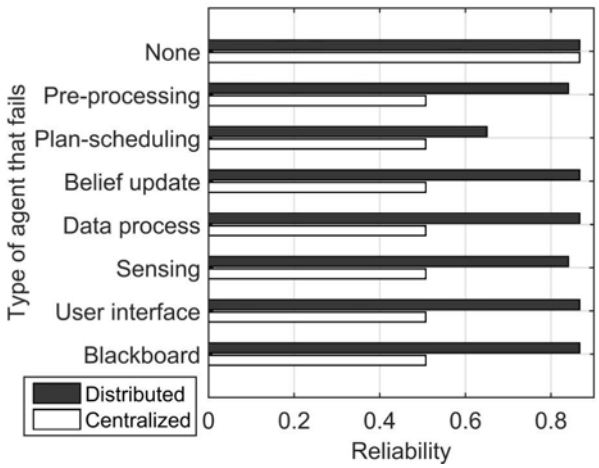


Fig. 15. The reliability of the system when each agent fails

lated and ranked, in order to introduce redundancy for the important ones to improve the reliability of the whole satellite system [11].

In above, the proposed online scheduling algorithm and the distributed multi-agent blackboard structure are efficient and effective for the autonomous satellite.

6. Conclusions

With the wide application of satellite technology, the traditional satellite controlling mode becomes unable to meet requirements of new complex tasks. In order to solve the problems of rapid response to emergencies, the environment uncertainties and the satellite faults, the new-generation autonomous satellite has become the focus of the satellite development.

In this paper, an autonomous multi-sensor satellite system based on the distributed multi-agent blackboard model is designed. Seven agents with different functions are designed. The agents interact with each other through the blackboard. The blackboard stores the satellite status information, task queue and action sequence, which are regarded as the belief set, desire set and intention set respectively. An adaptive rule-based heuristic scheduling algorithm is used to schedule the tasks. Four different operators are designed. And a fast forward search planning algorithm based on STRIPS is designed. The prototype system is realized by Jade 4.0. Four kinds of events are designed to verify the system's autonomy. The simulation experiments verify that the system can deal with the four kinds of events without human intervention, which greatly improves the performance and reliability of the satellite. The performance of the proposed algorithms and the distributed multi-agent blackboard structure is proved.

This work proves the effectiveness of using multi-agent technology to build the autonomous satellite system and contributes the distributed multi-agent blackboard model, which shows a high reliability and performance for those dynamic, uncertain and complex tasks. This structure also has high reconfigurability and extendibility. The blackboard model structure can be used to establish the future autonomous satellite system structure. The next focus of this research is to build a multi-satellite system using multi-agent technology to deal with concurrent processing and cooperative behaviors of the satellite constellation.

Acknowledgement

This paper is supported by National Natural Science Foundation of China (Grant No. 71501180 and 71501179) and the Foundation for the Author of National Excellent Doctoral Dissertation of China (Grant No. 201492). Thanks are due to the referees for their valuable comments.

References

1. Armstrong M. Joint reliability importance of elements. IEEE Transactions on Reliability 1995; 44(3): 408-412, https://doi.org/10.1109/24.406574.
2. Birnbaum Z. On the importance of different components in a multicomponent system. In P. Krishnaiah (Ed.), Multivariate Analysus-II. New York: Academic Press 1969.
3. Borgonovo E, Apostolakis G. A new importance measure for risk-informed decision making. Reliability Engineering and System Safety 2001; 72: 193-212, https://doi.org/10.1016/S0951-8320(00)00108-3.

4. Caskurlu B, Mkrtchyan V, Perekh O, Subramani K. On Pratial Vertex Cover and Budgeted Maximum Coverage Problems in Bipartite Graphs. Theoretical Computer Science: 8th IFIP TC 1/WG 2.2 International Conference, TCS 2014, Rome, Italy: Springer 2014; 13-25.
5. Curtis D E, Pemmaraju S V, Polgreen P. Budgeted Maximum Coverage with Overlapping Costs: Monitoring the Emerging Infections Network. 2010 Proceedings of the Twelfth Workshop on Algorithm Engineering and Experiments (ALENEX). Society for Industrial and Applied Mathematics 2010.
6. Du D, Ko K, Hu X. Design and analysis of appromaxition algorithms. Springer Optimization and Its Applications 2012, https://doi.org/10.1007/978-1-4614-1701-9.
7. Ericson II C A. Hazard Analysis technique for System Safety. New Jersey: John Wiley & Sons 2015
8. Espitrity J, Coit D, Prakash U. Component criticalty importance measures for the power industry. Electric Power Systems Research 2007; 407-420, https://doi.org/10.1016/j.epsr.2006.04.003.
9. GLPK (GNU Linear Programming Kit) From: www.gnu.org/software/glpk
10. Gupta S, Bachttacharya J, Barabady J, Kumar U. Cost-effective importance measure: A new approach for resource prioritization in a production plant. International Journal of Quality & Realibility Management 2013; 30 (4): 379-386, https://doi.org/10.1108/02656711311308376.
11. Khuller S, Moss A, Naor J. The budgeted maximum coverage problem. Information Processing Letters 1999; 70 (1): 39-45, https://doi.org/10.1016/S0020-0190(99)00031-9.
12. Kuo W, Zhu X. Importance measures in reliability, risk and optimization. Chichester: John Whiley & Sons 2012, https://doi.org/10.1002/9781118314593.
14. Rauzy A. A Benchmark of Boolean Formulae. From http://iml.univmrs.fr/~arauzy/aralia/ benchmark.htm
15. Revelle C S. A bibliography for some fundamental problem categories in discrete location science. European Journal of Operational Research, 2008; 184 (3): 817-848, https://doi.org/10.1016/j.ejor.2006.12.044.
16. Van der Borst M, Schoonakker H. An overview of PSA importance measures. Reliability Engineering and System Safety, 2001 72 (3): 241-245, https://doi.org/10.1016/S0951-8320(01)00007-2.
17. van Heuven van Staereling I, de Keijzer B, Schafer G. The Ground-Set-Cost Budgeted Maximum Coverage Problem. 41st International Symposium on Mathematical Foundations of Computer Science (MFCS 2016). Dagstuhl Research Online Publication Server, 2016
18. Vaurio J. Ideas and developments in importance measures and fault-tree techiques for reliability and risk analysis. Reliability Engineering and System Safety, 2010; 95: 99-107, https://doi.org/10.1016/j.ress.2009.08.006.
19. Vesely W, Davis T, Denning R, Saltos N. Measures of risk importance and their applications. Columbus: Battelle Columbus Labs, OH (USA), 1983, https://doi.org/10.2172/5786790.
20. Wu S. Joint importance of multistate system. Computers & Industrial Engineering, 2005; 49: 63-67, https://doi.org/10.1016/j.cie.2005.02.001.
21. Wu S, Coolen F. A cost-based importance measure for system components: An extension of the Birnbaum importance. European Journal of Operational Research, 2013; 189-195, https://doi.org/10.1016/j.ejor.2012.09.034.
22. Zafiropoulos E P, Dialynas N E. Methodology for the optimal component selection of electronic devices under reliability and cost constraints. Quality and Reliability Engineering International, 2007; 23 (8): 885-897, https://doi.org/10.1002/qre.850.
23. Zaitseva E, Levashenko V, Kostolny J. Application of logical differential calculus and binary decision diagram in importance analysis. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17 (3): 379-388, https://doi.org/10.17531/ein.2015.3.8.
24. Zio E. Risk importance measures. In H. Pham (Ed.), Safety importance measures and its applications. London: Springer, 2011; 151-196, https://doi.org/10.1007/978-0-85729-470-8_6.
25. Zio E, Podofilini L. Accouniting for components interactions in the differential importance measure. Reliability Engineering and System Safety, 2006; 91: 1163-1174, https://doi.org/10.1016/j.ress.2005.11.044.
26. Zio E, Podofilini L. The use of importance measures for the optimization of multi-state systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2006; 2: 33-36.

Lei HE
Guoliang LI
Lining XING
Yingwu CHEN
College of Information System and Management
National University of Defense Technology
No.47 Yanwachi Street, Changsha, China, 410073

E-mails: helei@nudt.edu.cn, worldchinali@126.com, xing2999@qq.com, ywchen@nudt.edu.cn

Article citation info:
GUTTEN M, KORENCIAK D, KUCERA M, SEBOK M, OPIELAK M, ZUKOWSKI P, KOLTUNOWICZ T. Maintenance diagnostics of transformers considering the influence of short-circuit currents during operation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 459–466, http://dx.doi.org/10.17531/ein.2017.3.17.

Miroslav GUTTEN
Daniel KORENCIAK
Matej KUCERA
Milan SEBOK
Marek OPIELAK
Pawel ZUKOWSKI
Tomasz N. KOLTUNOWICZ

MAINTENANCE DIAGNOSTICS OF TRANSFORMERS CONSIDERING THE INFLUENCE OF SHORT-CIRCUIT CURRENTS DURING OPERATION

EKSPLOATACYJNE DIAGNOSTYKI TRANSFORMATORÓW UWZGLĘDNIAJĄCE PRĄDY ZWARCIA PODCZAS PRACY

Article presents theoretical, simulation and experimental analyses of possible effect of short-circuit forces on the transformer windings. The first part of the article is focused to the establishment and activity radial and axial forces during short circuit. It shows dimensions, direction and of course caused mechanical stress. Equation shows basic dependencies of these mechanical forces created in the transformer windings. The last part of the article is focused on the simulation method which shows the mechanical stress caused by the short-circuit currents on transformer. The paper presents experimental methods of diagnostics for analysis of the short circuit on transformer windings.

***Keywords**: transformer, diagnostic, diagnostic system, transformer windings, frequency method.*

Artykuł przedstawia teoretyczną, symulacyjną i doświadczalną analizę ewentualnych wpływów prądów zwarciowych na uzwojenia transformatorów. Pierwsza część artykułu koncentruje się na powstaniu i działaniu promieniowych i osiowych sił podczas zwarcia. Przedstawia rozmiary, kierunek i oczywiście powstałe naprężenia mechaniczne. Równanie pokazuje podstawowe zależności sił mechanicznych powstałych w uzwojeniach transformatora. Ostatnia część pracy dotyczy metody symulacji, która przedstawia naprężenia mechaniczne spowodowane przez prądy zwarciowe na transformatorze. Artykuł przedstawia eksperymentalne metody diagnostyki do analizy wpływu zwarcia na uzwojenia transformatora.

***Słowa kluczowe**: transformator, diagnostyka, diagnostyczny system, uzwojenia transformatora, metoda częstotliwościowa.*

1. Introduction

The worst mechanical effects acting on the transformer are under the sudden short-circuit. Since the currents flows through the windings in the case of short-circuit are enormously, forces generated by these currents are also large and in many cases exceed the value of the maximum mechanical strength of windings or other parts of the transformer. Because of the transient and dynamics entity of this phenomenon the solution is very difficult.

It is therefore necessary, in the absence of scientific and research potential in distribution and transmission organizations (e.g. power plants, heating plants), to achieve the objectives of the proposed activities, i.e. in-depth analysis of undesirable impacts of short-circuit currents on the state of transformers, design of modern methodology of measurements and its verification, and design of a new integrated system of diagnostics with the possibility of analysis and classification of possible failures by short-circuit impact on power transformers.

Except the winding faults (inter-turn short-circuit, short time connection with the tank) it could lead to the tank damage because of the arc pressure. The direct cause of the acting forces on windings is the action of the magnetic field with current paths. In the case of the transformer it is the field of the leakage flux. Normal conditions, when the currents in the transformer do not reach the rated ones, the forces acting on the windings are generally small. On the contrary, the short-circuit currents reach the values which are multiple of the rated ones; these forces could be dangerous for the windings and fixing construction.

Forces in transformer windings can be radial (transverse) and axial (lengthwise), since these two components of the force can be calculated and analyzed independently. Also, the two components have influence on different parts of the total transformer and it is necessary to obtain the two components for design purposes. The nomenclature axial and radial is applicable to concentric wound core type transformer.

Action of the forces on the windings could be separated to the radial and axial forces (Fig. 1). Both of these forces exist not only in the short-circuit conditions but also in the rated condition [3].

2. Theory of radial and axial forces

Radial forces are generated by the interaction of the current and the axial component of the leakage flux density. They tend to squeeze the inner winding and burst the outer winding. The calculation of radi-

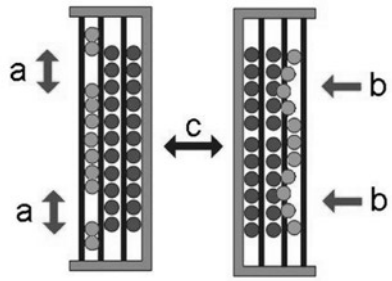


Fig. 1. Forces affect the winding of power transformer during short-circuit: a – axial effects, b – radial effects, c –winding in normal state

al forces does not at present much difficult since the axial component of the leakage flux density is calculable fairly accurately.

Calculation of the stresses due to this force acting on the conductors is more complex, especially for inner windings. The compressive strength of winding is influenced by the radial thickness of conductors, its work hardening strength, number of blocks per circle and winding dimensions. Similarly, the outer winding experiences outward radial force which results in tensile stress. The tensile strength of winding depends on the work hardening strength of the conductors.

Radial forces are as a result of an effect of the longitudinal field – parallel to the axis of winding of the transformer. These forces stretched the outdoor windings and compress the internal winding, thereby increase the air space between windings. The sum of the radial forces is active to additional space between the windings – the outdoor primary winding has been moved to the outside space [3].

Axial forces are those that act in the axial direction and are generated by the interaction of the current and radial component of the leakage flux density. These forces are flux result which has made by the short-circuited conductor.

These forces can deform the windings in the lengthwise direction, and their overall addition operates on the coil of clamping ring and other clamping elements. The computation of these forces causes big problems since the radial element of the leakage flux density is difficult to compute exactly. When ampere-turns are exactly aligned so that the leakage flux pattern is symmetrical, then the leakage field is axial over the major part of the coil height. But due to the flux lines dispersing in the radial direction in the vicinity of the winding ends, the axial flux density tends to decrease, and the resultant flux density at the ends can be resolved into the radial component causing axial forces. These axial forces are unequally distributed between the outer and inner winding, due to the presence of core.

The axial forces at the top and bottom are in opposite directions as the currents are in the same direction. In case the ampere-turns are perfectly balanced and the leakage flux pattern is symmetrical, the resultant force on the winding would be zero. Any axial displacement between the magnetic centres of HV and LV windings will result in forces, which will increase the displacement [5, 10].

2.1. Forces calculation

Based on theoretical analysis and model measurements it is needed to create a computing environment short-circuit forces. Based on this diagnostic approach it will be necessary to determine the possible effect of the devices and the insulation and mechanical condition of transformers.

The basic equation for the calculation of any electromagnetic force is:

$$\vec{F} = I\left(\vec{l} \times \vec{B}\right), \tag{1}$$

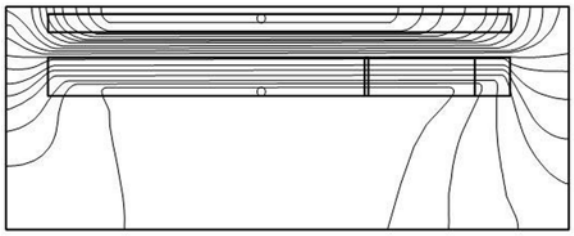


Fig. 2. Typical leakage flux pattern in a two winding transformer

where $\vec{B}$ is flux density due to leakage flux at mean radial depth of winding, $\vec{l}$ is the conductor length and I is the current during short-circuit. An examination of equation (1) shows that, since $\vec{F}$ is the cross-product of I and $\vec{B}$, a radial flux will give rise to axial force while an axial flux will give rise to radial force. The right side of equation (1) involves only three quantities, I, $\vec{B}$, and $\vec{l}$. Of these, the current I and the length $\vec{l}$ are known exactly. The calculation of the transverse and lengthwise components of the leakage flux density is the most difficult task and forms the crux of the problem.

2.2. Calculation of the radial forces

A simple formula for the average radial force can be derived by evaluating the axial component of the leakage flux density and applying equation (1). An assumption is made that all the leakage flux is axial only and passes between top and bottom yokes in straight lines. This assumption is also made in calculation of leakage reactance and will lead to a slightly higher value of radial force. The calculation of flux density of radial depth on transformer windings is given by:

$$B = \frac{1}{2} \cdot \ {}_0 \cdot \frac{N \cdot i_{\mathrm{p}}}{L_{\mathrm{u}}}. \tag{2}$$

The average value of the radial force after the substituting equation (2) to equation (1) given by:

$$F_{\mathrm{R}} = \frac{1}{2} \cdot \mu_0 \cdot \frac{\left(N \cdot i_{\mathrm{p}}\right)^2}{L_{\mathrm{u}}} \cdot \pi \cdot D_{\mathrm{m}}, \tag{3}$$

where μ_0 is a permeability of vacuum, N is number of turns, i_{p} is an asymmetric peak of the short-circuit current, L_{u} is height of the winding and D_{m} is a mean diameter of main inter-coil duct.

Due to the assumption that all the flux is axial in nature, equation (3) will result in a force larger than that due to any other formula. It is used in all stress calculations and in the design of windings.

The maximum force per turn will occur at the inner turn of innermost layer of the outer winding and the outer turn of the inner winding and will be given by:

$$F_{\mathrm{Rmax}} = \frac{F_{\mathrm{R}}}{N}. \tag{4}$$

If the turn experiencing this force has small radial dimension and is unsupported by any adjacent turns, the stress calculations are done using equation (4). Otherwise the average force of equation (3) is used for stress calculations. Graphical interpretation of the radial forces

acting on windings with interpretation of the symbols from equations (2) and (3) is shown in the Fig. 3.

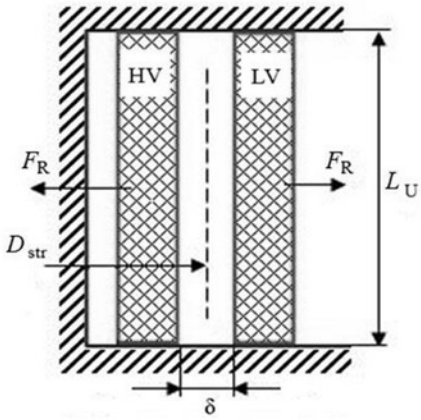


Fig. 3. Acting of the radial forces

2.3. Calculation of the axial forces

The electromagnetic forces arising due to a short circuit are oscillatory in nature, and act on an electric system immersed in oil and consisting of winding conductors, insulation components and clamping structure. Such forces, dynamically transmitted to various parts such as conductors, end supports, press plate, and clamps may be quite different, both in magnitude and in wave-shape from the internally generated electromagnetic forces depending on the relationship between the excitation frequency and the resonant frequency of the system.

When evaluating axial forces, winding misalignments, caused by workshop tolerances need to be considered. The design force calculations are performed both, for the symmetrical configuration and the displaced configuration of windings. In case of a symmetrical winding arrangement in axial direction having uniform current distribution, there is no resulting force against yoke, and the winding tends to be compressed in the axial direction only. Different yoke distances and tapping in the main winding and uneven current distribution in the axial direction can cause the force integral to reach a final value greater than zero. Two cases may occur. In the first case, the residual force is towards the bottom yoke. In the second case has a negative value at the bottom yoke, thus resulting residual force towards to top yoke.

While a simple formula could be obtained for radial forces, such is not the case for axial forces because must be taken into account the value of ampere-turns causing the radial leakage flux. An approximation to the ampere-turns causing radial leakage flux is the residual ampere-turns caused due to tapings, gaps in windings and unequal winding heights. This is only an approximation, since even perfectly

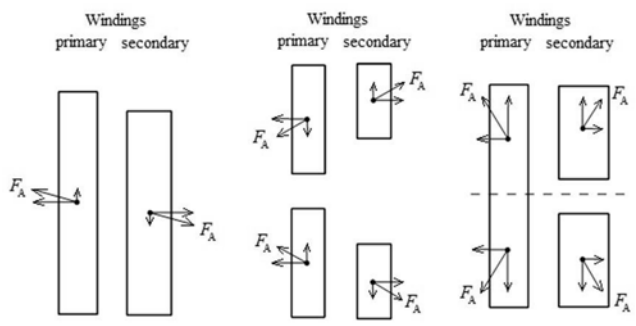


Fig. 4. Unbalanced windings causing axial forces

balanced windings will still have fringing of flux at the coil ends, giving rise to radial component of flux. The approximation is on the conservative side, since the peak value of the residual ampere-turns are used for calculations. Fig. 4 shows examples of unbalanced windings causing radial forces.

The flux lines are influenced by the core legs, yokes, the clamping structure, the tank walls and all other magnetic structures within the tank, and follow quite complex paths for closure. An exact evaluation of the effective length of path is extremely difficult, if not impossible.

For an acquisition of the empirical formulas various attempts were made. Hrabovcová V, et al. [1], Fergestad [2] and Knaack [8] proposed various methods for effective length evaluation but the simplest of all these approaches is that of Waters [12] which considering that there is an unbalance of ampere-turns.

The mean radial flux is given by:

$$B_R = \frac{1}{2} \cdot \mu_0 \cdot a \cdot \frac{N \cdot i_p}{l_{eff}}, \tag{5}$$

where a is per unit unbalance between windings and l_{eff} is the active length of track of short-circuit current.

Substituting (5) to the equation (1) and the follow fitting could be reached formula for axial force calculation:

$$F_A = \frac{1}{2} \cdot \mu_0 \cdot a \cdot \left(N \cdot i_p\right)^2 \cdot \frac{\pi \cdot D_m}{l_{eff}}, \tag{6}$$

where $\pi \cdot D_m / l_{eff}$ is the residual ampere-turns factor that values are stated in [12] for various ampere-turns configuration.

3. Simulation analysis

Simulation analysis involves calculation of the distribution transformer 22/0.4 kV model in a no load and short circuit conditions using COMSOL Multiphysics [6]. Second step is an approximation of the sinusoidal waveforms by calculating the values of the voltage and current using MATLAB Curve fitting tools. Next step consist the determination of the values of asymmetric peak of the short-circuit currents for short circuit directly behind the transformer, short circuit in the distribution box and short circuit on the end of the 1 km length cable. The final step is a calculation of the forces acting on the windings and their application to the modeled LV winding order to observe of their deformation and of the value of stress pressure. By applying the formulas for transformer measurements from [3] it is possible to calculate the values of equivalent circuit elements (Tab. 1).

Table 1. Values of the equivalent circuit elements

No load		Short-circuit	
$I_{\mu R}$	4.386 A	R_1	40.160 Ω
I_{feR}	1.326 A	R_2	0.013 Ω
$X_{\mu R}$	52.660 Ω	$X_{\sigma 1}$	91.016 Ω
R_{feR}	174.130 Ω	$X_{\sigma 1}$	0.030 Ω

3.1. Determination of the short-circuit currents

For the forces calculation is necessary to determine the value of the asymmetric peak of the short-circuit current. Program is structured for three different locations (points) of the short circuit (behind the transformer, in the distribution box, on the end of the cable). Fig. 5 shows connection diagram in MATLAB.

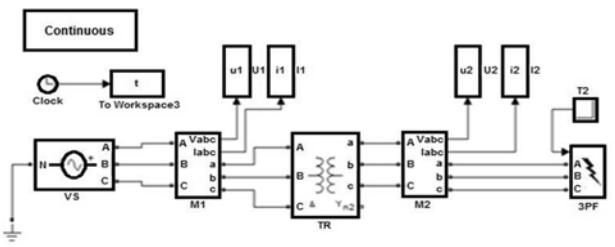


Fig. 5. Short circuit directly behind the transformer in MATLAB Simulink

Parameters of the circuit behind the transformer for different locations of short circuit are written in Tab. 2.

Table 2. Impedance parameters of the short circuit behind the transformer

Point of the short circuit	Circuit parameters
Behind the transformer	R=0 Ω, X=0 Ω
In the distribution box (15 m Cu 240 mm^2)	R=0.090 Ω/km, X=0.068 Ω/km
On the end of the cable (1 km Cu 120 mm^2)	R=0.181 Ω/km, X=0.069 Ω/km

Instantaneous value of short-circuit currents in LV windings is shown in Fig. 6. Values of the asymmetric peak of the short-circuit currents for different points of short circuit are shown in the Tab. 3.

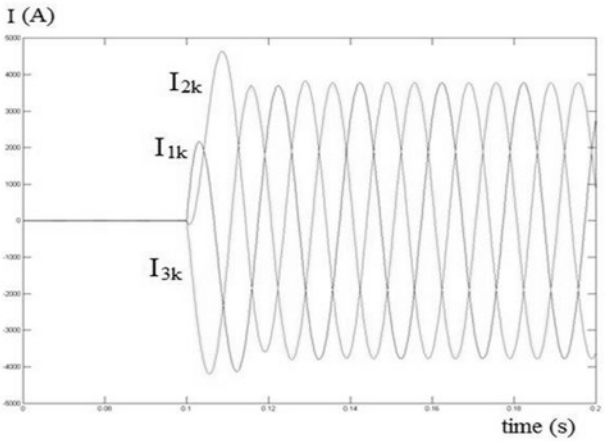


Fig. 6. Wave-shape of the short circuit behind the transformer

Since the short-circuit current has the largest value in phase B, the calculation of the forces is carried out with those most unfavorable values. An asymmetric peak of short-circuit current values in phase B for each point of the short circuit are shown in the Tab. 3.

3.2. Simulation in considering of the short-circuit currents

Calculation of the radial and axial forces for single layer of the LV winding is made by using the equations (3) and (6), where N is the turn number, i_p is an asymmetric peak of the short-circuit current, L_u is the height of the winding l_{eff} is active length of track of transverse leakage flux, a is per unit unbalance between windings and D_m is diameter of main inter-coil duct.

Values substituted to the equation for calculating of the forces are shown in the Tab. 4. The calculated values of the radial and axial forces for 5 % unbalance of the windings are shown in Tab. 5.

Fig. 7 shows the resulting pressure stress and permanent deformation of the winding due to radial force. Simulation of the radial force acting is calculated for the maximum value of the force (short circuit directly behind the transformer) applied to single layer of the LV winding.

Table 4. Values fitted to the equations (3) and (6)

Variable	Value
Conductor cross section	60 mm^2 (12 · 5)
Inner radius of the thread	141.4 mm ($100 \cdot \sqrt{2}$)
Number of turns in layer	32 (64 / 2)
Height of the winding	508.0 mm (12 · 32 + 4 · 31)
Effective path length of radial leakage flux	112.8 mm (0.222 · 508)
Mean diameter of thread	317.8 mm ($2 \cdot 100 \cdot \sqrt{2} + 35$)

Fig. 7. Permanent deformation and distribution of the pressure stress due to radial force

Fig. 8 shows the resulting pressure stress and permanent deformation of the winding due to axial force. Acting force is also proportional to the maximal current for the short circuit directly behind the transformer.

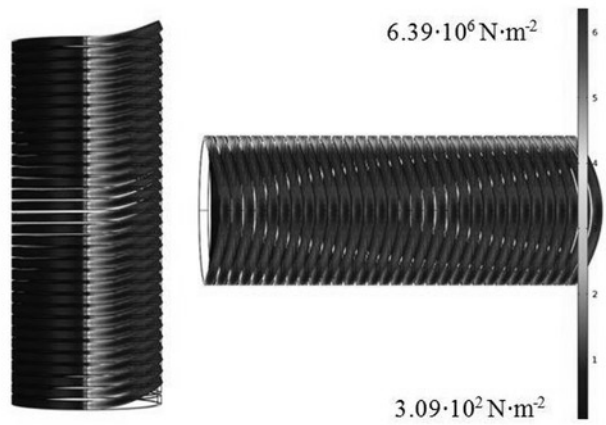


Fig. 8. Permanent deformation and distribution of the pressure stress due to axial force for 5 % per unit unbalance between windings

The dependence of the axial force on the unbalance between windings is shown on the Fig. 9. It also shows the necessity of precisely attachment of all windings in order to prevent of the per unit unbalance. Even at small values of unbalance greatly increases size of the axial force.

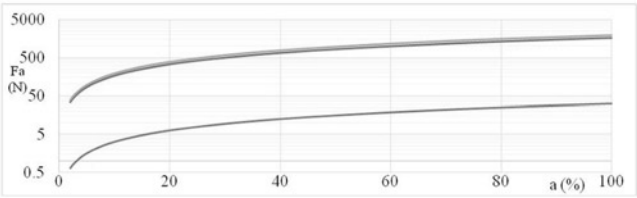


Fig. 9. Dependence of the axial force value on per unit unbalance between windings

4. The current experimental methods of diagnostics

The biggest problem of today's energy transmission and distribution works is that the measured data from the diagnostic measurements are not adequately studied, since there is no scientific and research base. Therefore, to assess technical condition of equipment, for example, in terms of their residual lifetime is a major problem in deciding on the early decommissioning. In an unpredictable moment, a failure may occur, and the consequence may be the power failure for a long time. In an unpredictable moment a failure may occur, and as a consequence the power failure for a long time. Even when using expensive commercial measuring devices, it is necessary to deeply analyze the measured data in relation to knowledge of exposure to adverse effects on specific electrical equipment. Thus achieving the objectives of this project may help to clarify the adverse effects and propose new diagnostic procedures through which it can be detected forthcoming failure of the device, and to suggest certain steps in advance (e.g. the early exchange for a new device) [11].

With regard to the effects of short-circuit currents, over-currents and other phenomena which damage winding and magnetic circuit, the following measurements will be realized at disconnected transformers:

- measurement of frequency characteristics by the SFRA method (Sweep Frequency Response Analysis),
- measurement of time response of windings by the high voltage impulse source (impact test),
- measurement of parameters of windings at short circuited state,
- combination of diagnostic methods according to the proposed measurement procedures.

The use of different diagnostic methods and measurement procedures enables to make measurements also at connected transformers. These are thermography measurements, and monitoring of changes in electromagnetic radiation of transformers and their influence on measure of interference with other devices, such as telecommunication or data devices.

At failure, e. g. as a result of shift or inter-turn short-circuit at the winding of a transformer, the electromagnetic interference with other electrical apparatuses may change.

That is why the following measurements at connected transformers will be realized with the use of experimental technique:

- thermography measurement,
- measurement with the regard to electromagnetic interference of distribution apparatuses,
- additional measurements, or their combination, according to the proposed measurement procedures.

Because of these measurements at transformers we can detect the effects of short-circuit currents, over-currents and other phenomena, which damage winding and magnetic circuit of the transformer, without its dismantling and subsequent investigating of failures, which is time-demanding [6].

Measurement of frequency characteristics by the SFRA method, measurement of time response of windings by the high-voltage impulse source and measurement of parameters of windings at short-circuited state belong to diagnostic methods of transformers without dismantling. During measurements there is no need of intervention to the construction of measured machine, and they are performed at disconnected transformer [13].

5. Experimental analysis of the distribution transformer

On the basis of applied algorithms and procedures realized in the first phase of solution in the paper, there were determined individual diagnostic methods and measuring techniques, for measurements on distribution transformer 22 kV / 0.4 kV with defect on the high voltage phase A (Fig. 10). Our task was to compare the sensitivity of the two methods – SFRA method and high voltage impulse source by impact test.

The SFRA method and by the measurement of time response method (impact test) enable to determine the time response or frequency response of characteristic quantities of transformers. Measurement of a response in the time domain means finding out of time course to a certain voltage pulse at the input of windings. Measurement of a response in the frequency domain consists finding out of amplitude (and phase in degrees) of a response to the harmonic voltage with a variable frequency at the input of winding. Whereas the response found out in the time domain is a record of time course of voltage, the response found out in the frequency domain is dependence of the amplitude response to frequency [9].

5.1. Analysis of the transformer by SFRA

The distribution transformer measurement applies for a setting up of the frequency scale at least 20 Hz to 2 MHz at output impedance 50 Ω and source voltage 0.2 – 20 V_{pp} using Megger FRAX 150 (Fig. 11). The measurements were performed in a no-load and short-circuited state according to established standards.

During no-load test is detected a mechanical state of tested winding and ferromagnetic core of transformer. The measured curves are typical for the most important information about changes in the transformer core, which are visible in low frequencies and higher frequencies refer to problems into winding movements or turn-to-turn fault regarding to short-circuit forces [4].

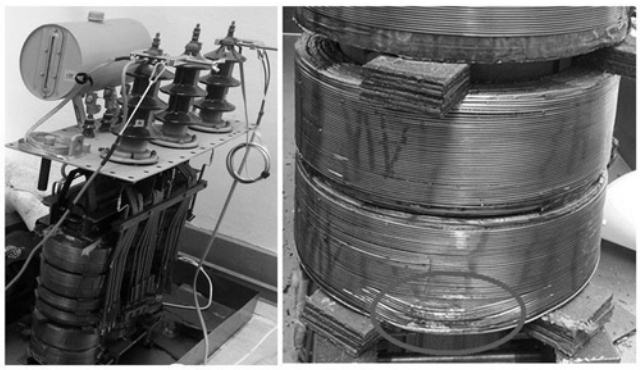


Fig. 10. Distribution transformer 22/0.4 kV under the test in HV laboratory with defect on the A phase

Difference between behaviours at low frequencies about to 1 kHz it can indicate problems due to movements of transformer core, because it developed its resistance component. Difference between behaviors at middle frequencies about to 100 kHz it can indicate problems due to deformation and damage of transformer core, because it developed inductive component of coils. At high frequencies about from 100 kHz it can indicate problems due to damage of winding

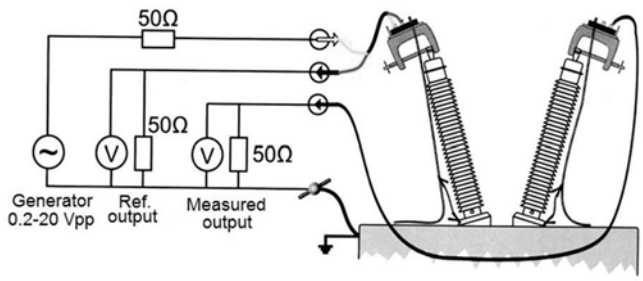


Fig. 11. SFRA test connection using FRAX 150

insulation or contact resistance of transformer taps, because it developed capacitance component.

During the short-circuit measurement mainly the winding state in primary or secondary part of transformer is detected. There tests notifies reliably of deformation of inner winding and its movement as a result effects of short-circuit forces.

The using of analysis of phase curve depending on frequency is suited for more complete evaluation of winding state. This analysis enables to assess the processes of winding deformation during the particular short-circuit effects.

In the Fig. 12 it can be seen the comparison of measurement curves of the impedance attenuation (dB) depending on frequency in a state of no-load and short-circuit at the connection primary phase A-C of measured transformer, and consequently in the Fig. 13 comparison of the curves of measurement the phase (in degrees) depending on the frequency of the same connection.

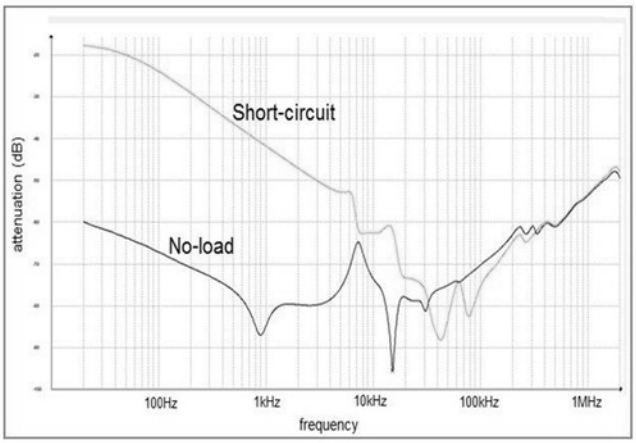


Fig. 12. Measurement of impedance attenuation

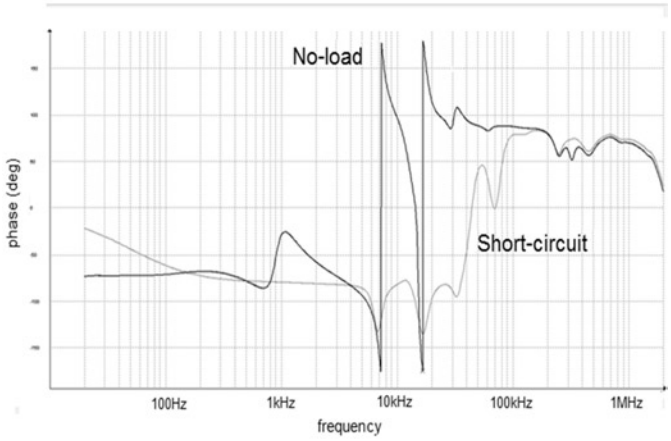


Fig. 13. Measurement of phase in degrees

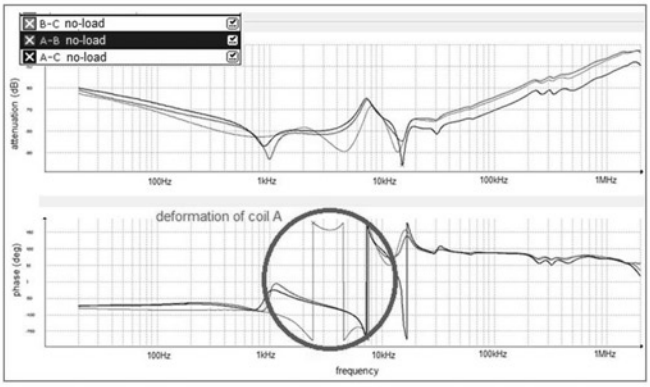


Fig. 14. Comparing the measurement of attenuation and phase in depending on the frequency in no-load test

According to comparison both figures it is highly visible difference of curves in depending phase from frequency in the Fig. 13, thereby increasing the analysis sensitivity of the transformer state.

Fig. 14 shows similar and significant difference. An inter-thread deformation on the coil A is most sensitive displayed by dependence of the phase on the frequency (about 1-10 kHz in lower figure). It should be noted that if the transformer windings are star connected, different curve between phases B – C means damage of other phase coil, thus coil A.

Analysis of the measured results is possible to carry out by normalized mathematical standard „DL/T 911-2004". This calculation method mathematically analyses differences between two curves (sequences) using relative factor R_{xy} in equations (7), (8) and (9).

Calculation of the two sequences standard variance:

$$D_x = \sum_{K=0}^{N-1}\left[X(k) - \frac{1}{N}\sum_{K=0}^{N-1}X(k)\right]^2 \quad D_Y = \frac{1}{N}\sum_{K=0}^{N-1}\left[Y(k) - \frac{1}{N}\sum_{K=0}^{N-1}Y(k)\right]^2. \quad (7)$$

Calculation of the two sequences covariance:

$$C_{XY} = \frac{1}{N}\sum_{K=0}^{N-1}\left[X(k) - \frac{1}{N}\sum_{K=0}^{N-1}X(k)\right]^2 \cdot \left[Y(k) - \frac{1}{N}\sum_{K=0}^{N-1}Y(k)\right]^2. \quad (8)$$

Calculation of the two sequences normalization factor covariance:

$$LR_{XY} = \frac{C_{XY}}{\sqrt{D_X D_Y}}. \quad (9)$$

Finally, from equation (9) is determined relative factor R_{xy}.

The result of the analyses in no-load conditions according equations (7) till (9) is value of the relative factor $R_{xy} = 0.35$ in the frequen-

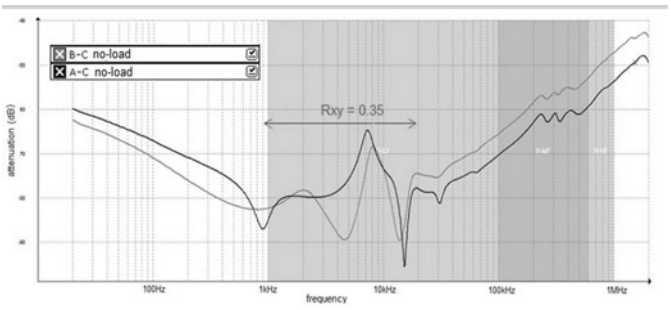


Fig. 15. Frequency analysis of deformation coil A between the winding-phase A-C and B-C

cy range from 0.9 kHz to 20 kHz (Fig. 15), where prevails value of the inductive part which is dependent on the geometry of the winding.

The finding above represents deformation anomaly due to inter-thread short-circuit, thus damage of the coil A and threat of the transformer operation (Fig. 10 – visual defect).

5.2. Analysis of the transformer by impact test

Impact test is commonly used for testing of the insulation among threads of coils or windings themselves and for searching of the attenuated winding sections of electrical machines. This test allows detecting early stages of the winding faults. Short time voltage pulses are applied to the coil in order to create a voltage gradient across the whole wire of the winding. This gradient produces an immediate voltage stress among the threads.

In the time intervals between pulses the coil responds by damping oscillations with sinusoidal shape. Each coil has unique nature of the respond which could be analysed by oscilloscope. Wave shape is affected by transient circuit dependent on the coil inductance and internal capacity of the pulse generator. When inter-thread occurs, an inductance decreases, thereby increases the frequency of the respond signal.

Fig. 16 shows a comparison of time-domain curves from impulse test measurement on the distribution transformer, where is possible to

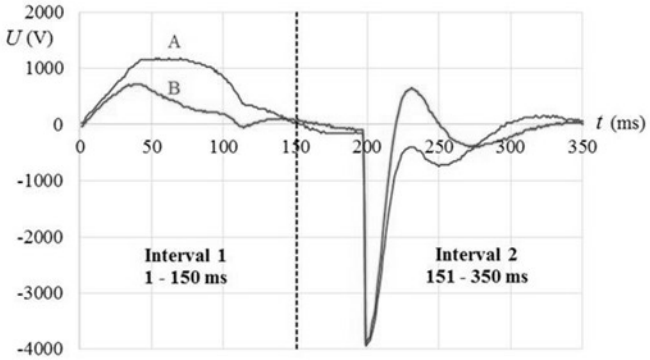


Fig. 16. Analysis of transformer coils by impact test

observe decreased time period from 9 ms to 12 ms caused by defect of the coil. Potential amplitude decrease is caused by the change of the circuit resistance and capacitance due to damaged conductor and winding insulation. The comparison is carried out on the windings of two phases, where phase A is affected by inter-thread short circuit and phase B is without fault.

In order to determine condition of windings it is necessary to understand that particular curves overlap when windings are identical and not damaged. Mutually shifted curves indicate damage on one of

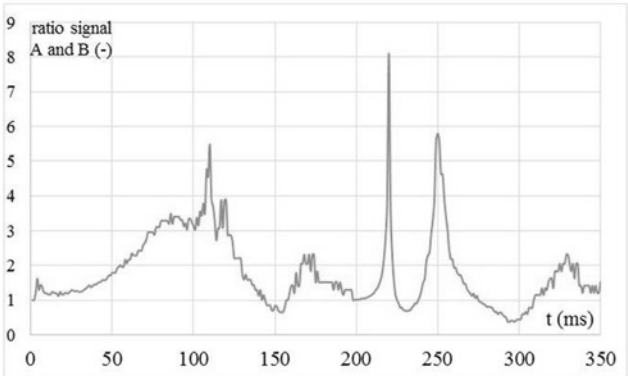


Fig. 17. Ratio of comparison signals A and B

the windings, therefore it is useful to analyse differences in curves at pulse courses.

Splitting impulses into intervals ensure more detailed analysis. Fig. 16 shows values of amplitude coefficient related to particular ratio in time intervals. So as to analyse changes in damaged windings, we compared the behaviour of ratio in time (see Fig. 17). In interval from 1 to 150 ms, maximum value of ratio is 5.47, in time interval from 151 to 350 ms, maximum of ratio is 8.11. These results show and prove that there is a short-cut as well as deformation on windings of A phases.

6. Conclusion

The article shows the importance of knowledge of simulation and theoretical analysis of the axial and radial forces, which cause irreversible deformation of the transformer windings. Above analysis shows that it is necessary to know the value of short-circuit current, because it represents a danger for the transformer.

Calculated values of short-circuit currents from the simulation in MATLAB were fit into the formulas for calculation of radial and axial forces. The calculated values were applied to the model of the winding by means COMSOL Multiphysics, which output of the simulation was to determine the size and distribution of compressive stresses.

The application of the short-circuit current is not possible on the real transformer. We can diagnose transformer by means of frequency SFRA (*Sweep Frequency Response Analysis*) method or time response of windings by the high voltage impulse source. Both methods, SFRA method and method of impact test, indicate a significant sensitivity to a relatively small inter-thread deformation. These methods allow clearly analysing of the short circuit impact in the transformer windings.

Although the relation between the response course and state of winding is clear, it is very complicated. It is impossible to expect, that from the found out difference in the course of response we can consider the concrete damage of the winding. As the result of measurement we can only observe, that the state of winding changed. Such result of using of mentioned above methods could be useful in decision making, whether it is necessary to open and examine the transformer.

In order to obtain frequency response curve, there is a procedure to recalculate time response of the impulse measurement to the frequency domain. However, without IFRA method (*Impulse Frequency Response Analysis*) equipment, the calculation is demanding and requires oscilloscope with large sampling frequency and memory. Information and outcomes mentioned in the paper are the basis for future investigation, which will focus on enlarging the knowledge of and determining clear relation among the frequency domain, time domain and condition of the transformer windings.

These diagnostic methods of transformers considering the influence of short-circuit currents during the operation should be carried out to increase the reliability in real trouble-free process.

Acknowledgment

This work was partly supported by the R&D operational program Centre of Excellence of Power Electronics Systems and Materials for their Components, No. OPVaV-2008/2.1/01-SORO, ITMS 26220120003 funded by the European Community and from the statute tasks of the Lublin University of Technology, at the Faculty of Electrical Engineering and Computer Science, 8620/E-361/S/2016 (S-28/E/2016), entitled "Researches of electrical, magnetic, thermal and mechanical properties of modern electrotechnical and electronic materials, including nanomaterials and electrical devices and their components, in order to determination of suitability for use in electrical engineering and to increase the efficiency of energy management".

References

1. Hrabovcová V, et al. Measuring and Modelling of Electrical Machines. Zilina: EDIS. 2009; 335. ISBN 978-80-8070-924-2.
2. Fergestad R. Electromagnetic Forces in Core-type Transformers with Concentric Windings. CIGRE. no. 114, Part II.
3. Gutten M. Analysis of short-circuit currents in electrical equipment. Zilina: EDIS. 2011; 103. ISBN 978-80-554-0433-2.
4. Chitaliya G. H., Joshi S. K. Finite Element Method for Designing and Analysis of the Transformer - A Retrospective. International Conference on Recent Trends in Power, Control and Instrumentation Engineering PCIE 2013. Hyderabad 2013, India.
5. Heatcote MJ. The J & P Transformer Book 13th edition. Chennai: ELSEVIER, 2007; 989. ISBN 978-0-7506-8164-3.
6. Gutten M. Diagnostic of distribution transformers by SFRA method. Przeglad Elektrotechniczny. 2007; 84(4): 144-146.
7. Glowacz A, Glowacz A, Glowacz Z. Recognition of Thermal Images of Direct Current Motor with Application of Area Perimeter Vector and Bayes Classifier. Measurement Science Review. 2015; 15(3): 119-126, https://doi.org/10.1515/msr-2015-0018.
8. Knaack W. The Mechanical Stressing of Transformer Windings Under Short-circuit. CIGRE. no. 135, Part II.
9. Kvasnička V, Procházka R, Velek J. Verification of method frequency characteristics in control room of distribution system Czech Republic. Conference Diagnostika 2005, Pilsen 2005, Czech Republic.
10. Sanka S. Transformers. 2nd edition. Bharat Heavy Electricals Limited Bhopal. New Dehli. 2003; 614. ISBN 978-0-17-048315-6.
11. Shayegani AA, Hassan O, Borsi H, Gockenbach E, Mosheni H. PDC Measurement Evaluation On Oil-Pressboard Samples. International Conference on Solid Dielectrics. 2014; 4: 50-62.
12. Heatcote M. J. J&P. Transformer Book 13th Edition. GB: ELSEVIER. 2007; 974 p. ISBN-13 978-0-7506-8164-3.
13. Werelius P, Öhlen M, Adeen L, Brynjebo E. Measurement Considerations using SFRA for Condition Assessment of Power Transformers. International Conference on Condition Monitoring and Diagnosis, Beijing, China, April 21-24, 2008, https://doi.org/10.1109/cmd.2008.4580428.

Miroslav GUTTEN
Daniel KORENCIAK
Matej KUCERA
Milan SEBOK
Faculty of Electrical Engineering
University of Zilina
1 Univerzitna Str., 01001 Zilina, Slovakia

Marek OPIELAK
Faculty of Mechanical Engineering
Lublin University of Technology
ul. Nadbystrzycka 36, 20-618 Lublin, Poland

Pawel ZUKOWSKI
Tomasz N. KOLTUNOWICZ
Faculty of Electrical Engineering and Computer Science
Lublin University of Technology
ul. Nadbystrzycka 38a, 20-618 Lublin, Poland

E-mails: miroslav.gutten@fel.uniza.sk, daniel.korenciak@fel.uniza.sk, matej.kucera@fel.uniza.sk, milan.sebok@fel.uniza.sk, m.opielak@pollub.pl, p.zhukowski@pollub.pl, t.koltunowicz@pollub.pl

Article citation info:
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, http://dx.doi.org/10.17531/ein.2017.3.18.

Rui PENG
Qingqing ZHAI

MODELING OF SOFTWARE FAULT DETECTION AND CORRECTION PROCESSES WITH FAULT DEPENDENCY

MODELOWANIE PROCESÓW WYKRYWANIA I KOREKCJI BŁĘDÓW OPROGRAMOWANIA Z ZAŁOŻENIEM WZAJEMNEJ ZALEŻNOŚCI BŁĘDÓW

Software reliability modeling has undergone a continuous evolution over the past three decades to adapt to various and ever-changing software testing environments. In existing models, immediate fault removal and fault independency are two basic and commonly used assumptions. Recently, models combining fault detection process (FDP) and fault correction process (FCP) were proposed to alleviate the immediate fault removal assumption. In this paper, we extend such a methodology by proposing a modeling framework for the FDP and FCP incorporating fault dependency. Faults are classified as leading faults and dependent faults and the FCPs for both types of faults are explicitly modeled. Several paired models considering different assumptions for debugging lags are proposed for the combined FDP and FCP. The applicability of the proposed models are illustrated using real testing data. In addition, the optimal software release policy under this framework is studied.

Keywords: *fault dependency, non-homogeneous Poisson process (NHPP), software reliability growth model (SRGM), software fault detection and correction processes.*

Modelowanie niezawodności oprogramowania w ciągu ostatnich trzech dekad ulegało ciągłej ewolucji, pozwalającej dostosować je do różnych, stale zmieniających się środowisk testowych. W przypadku istniejących modeli, dwoma podstawowymi i powszechnie stosowanymi założeniami jest natychmiastowe usunięcie błędu oraz brak zależności między błędami. Ostatnio, badacze zaproponowali modele, które łagodzą pierwsze z tych założeń, łącząc proces wykrywania błędów (FDP) z procesem ich korekcji (FCP). W niniejszym artykule, rozszerzono tę metodologię, proponując paradygmat modelowania dla zintegrowanych procesów FDP i FCP uwzględniający zależności między błędami. W paradygmacie tym, błędy klasyfikuje się jako błędy nadrzędne i błędy zależne, a procesy FCP dla obu typów błędów są modelowane oddzielnie. Zaproponowano kilka połączonych w pary modeli rozważających różne założenia dotyczące opóźnień debugowania w procesach łączących detekcję i korekcję błędów. Możliwość zastosowania proponowanych modeli przedstawiono na przykładzie rzeczywistych danych testowych. Dodatkowo badano optymalną politykę aktualizacji oprogramowania, jaką można prowadzić w ramach proponowanego paradygmatu.

Słowa kluczowe: *zależność błędów, niejednorodny proces Poissona, model wzrostu niezawodności oprogramowania, procesy detekcji i korekcji błędów oprogramowania*

1. Introduction

Software today plays important roles in almost every section of our society, and the software reliability has been a major concern in many integrated systems [3]. With continuous debugging, analysis and correction, the software reliability will grow gradually with testing [33]. During the past three decades, numerous software reliability growth models (SRGMs) have been proposed [2, 7, 24, 26, 35, 40, 41]. Among these models, Non-homogeneous Poisson Process (NHPP) models are the most commonly accepted [20, 30, 36, 39, 50]. Although NHPP models are mathematically tractable, they are developed under some strong assumptions on the software testing process. Specifically, NHPP models assume immediate fault removal and fault independency. To adapt to different practical software testing environments, generalizations of traditional models by relaxing the assumptions have been proposed [5, 9, 17, 23, 28, 29].

In practical software testing, each detected fault has to be reported, diagnosed, removed and verified before it can be noted as corrected. Consequently, the time spent for fault correction activity is not negligible. In fact, this debugging lag can be an important element in making decisions [16, 49]. Therefore, it is necessary to incorporate the debugging lag into the modeling framework, *i.e.*, to model both the fault detection process (FDP) and fault correction process (FCP). The idea of modeling FCP was first proposed in Schneidewind [34], where a constant lag was used to model the FCP after fault detection. Clearly, the constant correction time assumption is restrictive for various types of faults and different correction profiles. For instance, data collected from practical testing projects show that the correction time can be fitted by the exponential and log-normal distributions [27]. In addition, the correction time may show a growing trend during the whole testing cycle, as later detected faults can be more difficult to correct. Some extensions were made in Lo and Huang [25] and Xie, et al. [44] by incorporating other assumptions of debugging delay. Hu, et al. [8] studied a data-driven artificial neural network model for the prediction of FDP and FCP. [37] used the fault detection/correction profile to quantify the maintainability of software. Some paired FDP and FCP models were proposed in Peng, et al. [31], where testing effort function and fault introduction were included.

Traditional NHPP models assume the statistical independency between successive software failures. Actually, it can hardly be true in practice, as some faults are not detectable until some other fault has been corrected because of logical dependency. Moreover, the common practice of mixing testing strategies can lead to the dependency of failures [6]. With a failure detected, there is a higher chance for another related failure or a cluster of failures to occur in the near future. From this point of view, faults can be classified into mutually independent and dependent types with respect to path-based testing approach. This

issue was addressed in [18], where an extended NHPP SRGM was proposed. Huang and Lin [11] studied the fault detection & correction process considering both fault dependency and debugging lags. Yang, et al. [46] discussed the statistical inference of the software reliability model with fault dependency. However, most of the studies only focus on the FDP, and only the FDP data are used for model parameters estimation. As a result, the collected information from FCP is neglected, which can lead to deficiency in model estimation.

To remedy the problem, we incorporate the fault dependency into the paired FDP and FCP model. Instead of assuming a single type of fault, this study classifies the faults in the testing process into leading faults and dependent faults.The leading faults occurs independently following an NHPP, while the dependent faults are only detectable after the related leading faults being corrected. Different from Huang and Lin [11] which modeled the FDP and the FCP as a single, synthesized fault detection & correction process, we model the FDP and FCP for the leading faults and the dependent faults separately. Subsequently, the FDP&FCP model for the aggregated, observable faults can be readily obtained. With different formulation of debug delays, we can derive various FDP&FCP models. Hence, the proposed models admit a wide applicability that can account for different software reliability growth schemes.

The rest of this paper is organized as follows. Section 2 formulates the general modeling framework of paired FDP and FCP with the incorporation of fault dependency. In Section 3, special paired FDP and FCP models are derived based on different assumptions for debugging lags. In Section 4, the proposed faults are fitted to two real datasets to illustrate the application.Section 5 derives the optimal software release policy under the proposed framework. The conclusion is given in Section 6.

Notation

a	The total number of faults in the software
a_1	The number of leading faults in the software
a_2	The number of dependent faults in the software
p	The ratio of the number of leading faults to the total number of faults
$b(t)$	Fault detection rate function at time t
b	Constant fault detection rate
c	Constant fault correction rate
$\delta(t)$	The time required to correct a fault which is finally corrected at time t
$m_d(t)$	Expected number of faults detected up to time t
$m_r(t)$	Expected number of faults removed up to time t
$m_{d1}(t)$	Expected number of leading faults detected up to time t
$m_{r1}(t)$	Expected number of leading faults removed up to time t
$m_{d2}(t)$	Expected number of dependent faults detected up to time t
$m_{r2}(t)$	Expected number of dependent faults removed up to time t
$\lambda_d(t)$	The intensity function of fault detection process
$\lambda_r(t)$	The intensity function of fault correction process
$\lambda_{d1}(t)$	The intensity function offault detection process for leading faults
$\lambda_{r1}(t)$	The intensity function of fault correction process for leading faults
$\lambda_{d2}(t)$	The intensity function offault detection process for dependent faults
$\lambda_{r2}(t)$	The intensity function of fault correction process for dependent faults

2. The general framework

In this study, we formulate the fault-oriented software testing process as a paired fault detection and correction process. During the test, a fault can only be corrected after being detected. For the faults embedded in the software system, they can be categorized into leading faults and dependent ones. The faults that can be detected and corrected independently are defined as leading faults or independent faults. Other faults that remain undetectable until the corresponding leading faults are removed are defined as dependent faults. Fig.1 illustrates the relationship between leading faults and dependent faults.

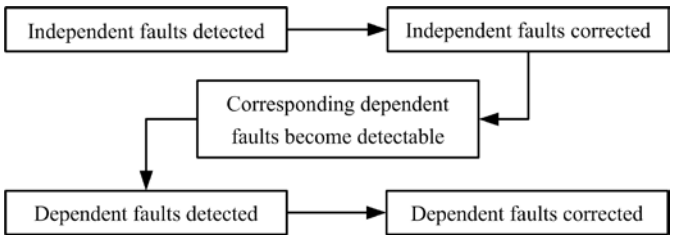


Fig. 1. Relationship of leading faults and dependent faults

Suppose the leading faults are detected and corrected independently. Then, for the leading faults, their detection (FDP_L) and correction process (FCP_L) can be modeled by NHPP models, as in Xie, et al. [44]. For the dependent faults, their detection process (FDP_D) can be modeled as a delayed process of FCP_L, considering that they are only detectable after the corresponding leading faults are corrected. Consequently, the correction process for dependent faults (FCP_D) can be modeled as a delayed process of FDP_D. The modeling framework is characterized by the mean value function for each sub-process.

2.1. Modeling FDP_L

We assume that FDP_L follows a NHPP, and the expected number of leading fault detected during $(t, t+\Delta t]$ is proportional to the number of undetected leading faults at time t. Thus we have:

$$\frac{dm_{d1}(t)}{dt} = b(t)\left(a_1 - m_{d1}(t)\right), \tag{1}$$

where $b(t)$ is the fault detection rate at time t and a_1 is number of leading faults at the beginning. With the initial condition $m_{d1}(t)=0$, it can be derived from (1) that:

$$m_{d1}(t) = a_1\left(1 - \exp\left\{-\int_0^t b(s)\,ds\right\}\right). \tag{2}$$

Different $m_{d1}(t)$ can be obtained based on different $b(t)$. Specially, when $b(t)$ is a constant, we have:

$$m_{d1}(t) = a_1\left(1 - e^{-bt}\right), \tag{3}$$

which is the G-O model [4]. When $b(t) = \frac{b^2 t}{1+bt}$, we have:

$$m_{d1}(t) = a_1\left(1 - (1+bt)e^{-bt}\right), \tag{4}$$

which has the same form as the Yamada delayed-S-shaped model [45].

2.2. Modeling FCP_L

FCP_L can be regarded as a delayed process of FDP_L and different models can be used to accommodate the debugging delay. Xie, et al. [44] pointed out that debugging lags could be assumed constant, time dependent or random. If the debugging lag is not random, the FCP_L can be derived from FDP_L as $m_{r1}(t)=m_{d1}(t-\delta(t))$. Otherwise, we have $m_{r1}(t)=E_{\delta(t)}\left[m_{d1}(t-\delta(t))\right]$ if $\delta(t)$ is a random variable. In particular, if the debugging lag is assumed to be an exponentially distributed random variable, i.e., $\delta(t)\sim Exp(c)$, we have:

$$m_{r1}(t)=c\int_0^t m_{d1}(t-s)\exp\{-cs\}\,ds. \tag{5}$$

Taking the derivatives of both sides with respect to t, we can obtain that:

$$\lambda_{r1}(t)=c\left(m_{d1}(t)-m_{r1}(t)\right). \tag{6}$$

This implies that the expected number of faults corrected during $(t,t+\Delta t]$ is proportional to the number of detected but uncorrected faults at time t. We call c the fault correction rate.

2.3. Modeling FDP_D and FCP_D

For these dependent faults, they can only be detected after the corresponding leading faults are removed. Hence, the proportion of the detectable dependent faults in the dependent faults is equal to the proportion of the corrected leading faults in the leading faults. Suppose the number of dependent faults is a_2. Then, the expected number of detectable dependent faults is $a_2 m_{r1}(t)/a_1$ up to time t. Furthermore, because leading faults and dependent faults are detected under the same testing environment, it is reasonable to assume that the fault detection rate for dependent faults is the same as the fault detection rate for leading faults. Therefore:

$$\frac{dm_{d2}(t)}{dt}=b(t)\left(\frac{a_2 m_{r1}(t)}{a_1}-m_{d2}(t)\right). \tag{7}$$

With the initial condition $m_{d2}(0)=0$, we can derive from (7) based on $m_{r1}(t)$ and $b(t)$ that:

$$m_{d2}(t)=\frac{a_2}{a_1}m_{r1}(t)-\frac{a_2}{a_1}\exp\left\{-\int_0^t b(s)\,ds\right\}\int_0^t \lambda_{r1}(s)\exp\left\{\int_0^s b(u)\,du\right\}ds. \tag{8}$$

Particularly, when $b(t)=b$, we have:

$$m_{d2}(t)=\frac{a_2}{a_1}m_{r1}(t)-\frac{a_2}{a_1}\exp\{-bt\}\int_0^t \lambda_{r1}(s)\exp\{bs\}\,ds. \tag{9}$$

Based on the detection process of dependent faults, the corresponding correction process can be obtained as a delayed process as for leading faults. Thus, with different assumptions for the debugging delay, $m_{r2}(t)$ of FCP_D can be derived accordingly.

2.4. Combined models

With the FDP and FCP models for both kinds of faults, the aggregated model for the paired FDP&FCP can be readily obtained:

$$a=a_1+a_2, \tag{10}$$

$$m_d(t)=m_{d1}(t)+m_{d2}(t), \tag{11}$$

$$m_r(t)=m_{r1}(t)+m_{r2}(t). \tag{12}$$

3. Specific models for dependent FDP and FCP

In this section, we consider the widely-used constant fault detection rate function $b(t)$, i.e., $b(t)=b$ [10, 22]. In this case, we have $m_{d1}(t)=ap\left(1-e^{-bt}\right)$ from (3), where $p=a_1/a$ is the proportion of leading faults. As stated, different $m_{r1}(t)$ can be derived based on different assumptions on the debugging lag. Moreover, as long as $m_{r1}(t)$ being specified, $m_{d2}(t)$ can be obtained according to (9). In the following, we consider three different types of debugging lags, which have been observed from practical testing processes. Correspondingly, specific paired PDF&FCP models are derived.

3.1. Constant debugging lag

We first consider the case where the correction of each fault takes the same time, i.e., $\delta(t)=\delta$. Then, the FCP model of leading faults is:

$$m_{r1}(t)=\begin{cases}0, t<\delta\\ ap\left(1-e^{-b(t-\delta)}\right), t\ge\delta\end{cases}. \tag{13}$$

Consequently, the FDP model for dependent faults can be derived according to (9):

$$m_{d2}(t)=\begin{cases}0, t<\delta\\ a(1-p)\left(1-\left(1+b(t-\delta)\right)e^{-b(t-\delta)}\right), t\ge\delta\end{cases}. \tag{14}$$

Based on the FDP models of the leading faults and the dependent faults, $m_d(t)$ for the aggregated FDP is obtained as:

$$m_d(t)=\begin{cases}ap\left(1-e^{-bt}\right), t<\delta\\ ap\left(1-e^{-bt}\right)+a(1-p)\left(1-\left(1+b(t-\delta)\right)e^{-b(t-\delta)}\right), t\ge\delta\end{cases}. \tag{15}$$

Because the FCP models for both kinds of faults are modeled as delayed FDP, the aggregated FCP model is:

$$m_r(t)=\begin{cases}0, t<\delta\\ ap\left(1-e^{-b(t-\delta)}\right), \delta\le t<2\delta\\ ap\left(1-e^{-b(t-\delta)}\right)+a(1-p)\left(1-\left(1+b(t-2\delta)\right)e^{-b(t-2\delta)}\right), t\ge 2\delta\end{cases} \tag{16}$$

3.2. Time-dependent Debugging Lag

In practice, the faults discovered in the later phase of the testing process may be more difficult to correct. To model such a phenomenon, we assume the debugging lag is dependent on the testing time, $\delta(t)=\frac{\ln(1+\gamma t)}{b}$, where $0<\gamma<b$. Accordingly, the FCP models for the

two kinds of faults are $m_{ri}(t)=m_{di}\left(t-\frac{\ln(1+\gamma t)}{b}\right), i=1,2$. Under this assumption, we have:

$$m_{r1}(t)=m_{d1}\left(t-\frac{\ln(1+\gamma t)}{b}\right)=ap\left(1-(1+\gamma t)e^{-bt}\right), \quad (17)$$

which is a general form of the delayed NHPP model [45].

Based on (9) and (17), $m_{d2}(t)$ can be derived. Then, $m_d(t)$ for the aggregated FDP is obtained as:

$$m_d(t)=a\left(1-e^{-bt}\right)-a(1-p)\left(bt+\frac{b\gamma t^2}{2}\right)e^{-bt}. \quad (18)$$

Because $m_r(t)=m_d\left(t-\frac{\ln(1+\gamma t)}{b}\right)$, the model for the aggregated FCP can be derived as follows:

$$m_r(t)=a\left(1-(1+\gamma t)e^{-bt}\right)$$
$$-a(1-p)(1+\gamma t)\left(bt-(1+\gamma t)\ln(1+\gamma t)+\frac{b\gamma t^2}{2}+\frac{\gamma\ln^2(1+\gamma t)}{2b}\right)e^{-bt}. \quad (19)$$

3.3. Exponentially distributed random debugging lag

As obtained in Section 2.2, the number of faults corrected during time interval $(t,t+\Delta t]$ in this case is proportional to the number of detected but uncorrected faults at time t. Based on (5), $m_{r1}(t)$ can be obtained as:

$$m_{r1}(t)=\begin{cases} ap\left(1-(1+bt)e^{-bt}\right), c=b \\ ap\left(1+\frac{be^{-ct}-ce^{-bt}}{c-b}\right), c\neq b \end{cases}. \quad (20)$$

Then, $m_{d2}(t)$ can be derived based on $m_{r1}(t)$ according to (9). As the summation of $m_{d1}(t)$ and $m_{d2}(t)$, $m_d(t)$ for the aggregated FDP is thus obtained:

$$m_d(t)=\begin{cases} a\left(1-e^{-bt}\right)-a(1-p)\left(bt+\frac{b^2t^2}{2}\right)e^{-bt}, c=b \\ a\left(1-e^{-bt}\right)-a(1-p)\left(\frac{bcte^{-bt}}{c-b}+\frac{b^2\left(e^{-ct}-e^{-bt}\right)}{(c-b)^2}\right), c\neq b \end{cases} \quad (21)$$

From $m_{ri}(t)=c\int_0^t m_{di}(t-s)\exp\{-cs\}ds$, we note that $m_r(t)=c\int_0^t m_d(t-s)\exp\{-cs\}ds$ also holds. Therefore, $m_r(t)$ is readily obtained as:

$$m_r(t)=\begin{cases} a\left(1-(1+bt)e^{-bt}\right)-a(1-p)\left(\frac{b^2t^2}{2}+\frac{b^3t^3}{6}\right)e^{-bt}, c=b \\ a\left(1+\frac{be^{-ct}-ce^{-bt}}{c-b}\right)-\frac{abc(1-p)}{(c-b)^2}\left(\frac{(b+c)\left(e^{-ct}-e^{-bt}\right)}{c-b}+bte^{-ct}+cte^{-bt}\right), c\neq b \end{cases} \quad (22)$$

4. Numerical example

In this section, we illustrate the application of the proposed models to two real software testing datasets.

4.1. Description of the Datasets

The first dataset is from the System T1 data of the Rome Air Development Center (RADC) [27]. This dataset is widely used and it contains both fault detection data and fault correction data. The cumulative numbers of detected faults and corrected faults during the first 21 weeks are shown in Table 1. During the debugging, 300.1 hours of computer time were consumed and 136 faults were removed. The computer time spent in the testing process is used the time scale for the FDP and FCP.

Table 1. The dataset from System T1.

Weeks	Computer time (CPU hours)	Cumulative number of detected faults (m_d)	Cumulative number of corrected faults (m_r)
1	4	2	1
2	8.3	2	2
3	10.3	2	2
4	10.9	3	3
5	13.2	4	4
6	14.8	6	4
7	16.6	7	5
8	31.3	16	7
9	56.4	29	13
10	60.9	31	17
11	70.4	42	18
12	78.9	44	32
13	108.4	55	37
14	130.4	69	56
15	169.9	87	75
16	195.9	99	85
17	220.9	111	97
18	252.3	126	117
19	282.3	132	129
20	295.1	135	131
21	300.1	136	136

The second dataset is from the testing process of a middle-size software project [42, 44]. The cumulative numbers of detected faults and corrected faults during the first 17 weeks are listed in Table 2.

Table 2. The dataset from a middle-size software project.

Weeks	Cumulative number of detected faults (m_d)	Cumulative number of corrected faults (m_r)
1	12	3
2	23	3
3	43	12
4	64	32
5	84	53
6	97	78
7	109	89
8	111	98
9	112	107
10	114	109
11	116	113
12	123	120
13	126	125
14	128	127
15	132	127
16	141	135
17	144	143

4.2. Performance analysis

To illustrate our models, we consider the following three paired FDP&FCP models: (1) model with constant debugging lag (abbreviated as M1); (2) model with $\delta(t)=\frac{\ln(1+\gamma t)}{b}$ (abbreviated as M2); and (3) model with exponentially distributed debugging lag (abbreviated M3).

We note that the models proposed in Xie, et al. [44] are special cases of the proposed FCP&FDP models without considering the dependent faults. For comparison purpose, we also fit the data by the three simplified models of M1-M3 with p=1, which are abbreviated as M1', M2' and M3', respectively.

The six models are fitted to the two datasets by the least squares method. The least squares method minimizes the mean squared error (MSE) between the estimated cumulative numbers of detected and corrected faults and the actual cumulative numbers of detected and corrected faults. It is calculated as:

$$\mathrm{MSE}=\frac{1}{2}\left(\mathrm{MSE}_d+\mathrm{MSE}_r\right)=\frac{1}{2n}\sum_{i=1}^{n}\left[\left(m_d(t_i)-m_{d,i}\right)^2+\left(m_r(t_i)-m_{r,i}\right)^2\right], \quad (23)$$

where $m_{d,i}$, $m_{r,i}$ are the observed cumulative numbers of detected faults and corrected faults at time t_i,i=1,…,n. The estimated model parameters for dataset 1 is given in Table 3.

Table 3. The estimated model parameters for dataset 1.

Model	a	b	Remark	MSE
M1	199.27	0.00717	δ=24.78, p=0.3820	9.0114
M1'	507.47	0.00110	δ=25.71	10.8924
M2	182.23	0.00955	γ=0.1599, p=0.1374	15.5697
M2'	1737.64	0.000288	γ=0.0930	26.1383
M3	185.15	0.008456	c=0.03833, p=0.3265	7.8881
M3'	477.75	0.001177	c=0.03786	10.0985

As can be noticed from Table 1, the estimated parameter a (the total number of faults) in the three proposed models M1-M3 are close to each other. They are all close to 188, which is the number of detected faults after three years' testing, as reported in Kapur and Younes [18]. On the contrary, the models M1'-M3', which assume no dependent faults exist, produce quite large a. Therefore, ignoring the dependent faults in the model would result in incorrect total number of faults.

According to the MSE values and the point-wise squared error $\mathrm{MSE}_{d,i}$+$\mathrm{MSE}_{r,i}$ in Fig. 2, it shows that the paired FDP&FCP model with exponentially distributed debugging lag fits the dataset best. On the other hand, the model M1, which assumes constant debugging lag, also provides a competitive fit. The model assuming time-dependent debugging lags provides the least favorable fit. In fact, according to the estimated model M3, we can derive that the expected length of the debugging lag is $\frac{1}{c}=26.08$. This is close to the estimated debugging lag in M1. Thus, we can infer that there are significant debugging lags in the software testing process, and it takes about 25 hours for a detected fault to be corrected.

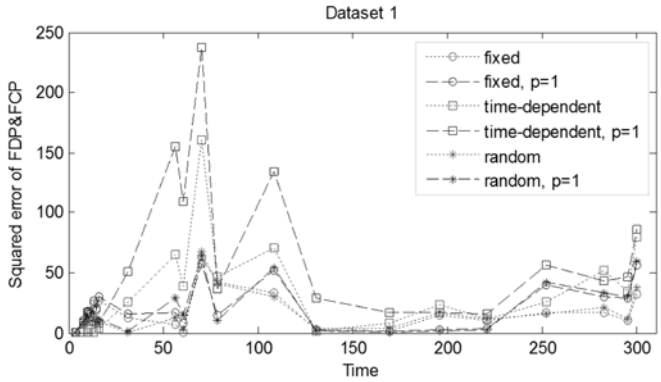


Fig. 2. Point-wise squared errors of the six fitted models for dataset 1.

The estimation results by the six models for dataset 2 are presented in Table 4. Analogous to the dataset 1, the proposed models considering both leading and dependent faults are superior to those

Table 4. The estimated model parameters for dataset 2.

Model	a	b	Remark	MSE
M1	144	0.3058	δ=1.51,p=0.474	39.5732
M1'	153.01	0.1487	δ=1.94	41.0015
M2	144	0.3938	γ=0.3112,p=0.0448	49.9352
M2'	168.36	0.1193	γ=0.2339	104.8889
M3	144	0.3354	c=0.7281,p=0.3551	47.0471
M3'	156.35	0.1404	c=0.5811	55.1920

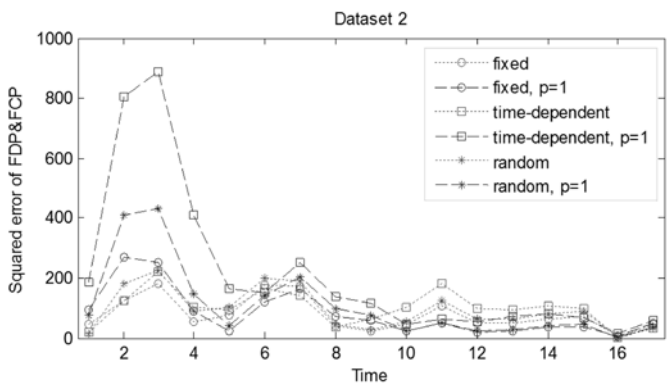


Fig. 3. Point-wise squared errors of the six fitted models for dataset 2

only considering leading faults. In the fitting procedure, we restrict the total number of faults a to be no smaller than the faults in the data. Therefore, we see that the estimated a are all equal to 144, which is the number of the total faults in dataset 2. Among the three models M1-M3, the constant debugging lag model provides the best fit. This can also be noted from the point-wise squared error in Fig. 3. This indicates the debugging lag is almost constant in the software testing process.

5. Software release policy

Based on a SRGM, useful information can be inferred to guide decision-making. For software projects, one critical decision is to determine the optimal release time [12]. Many studies have dealt with this problem [13, 19, 21, 32]; see Jain and Priya [14] and Boland and Chuív [1] for an overview. As cost and reliability requirements are of great concern, they are often used as objectives for optimizing the testing time and release policy [15, 38, 47, 48]. In this section, we study the optimal release policies based on the proposed models from the cost and reliability perspectives.

5.1. Software release policy based on reliability criterion

Software is usually released when a reliability target is achieved. It is reasonable to stop testing when a pre-specified proportion of faults are removed. We use T to denote the length of testing and consider the ratio of cumulative removed faults to the initial faults in the software system as the reliability criterion [11]:

$$R_1(T) = \frac{m_r(T)}{a}. \tag{24}$$

With a given reliability target R_1, the time to reach this reliability target is

$$T_1 = m_r^{-1}(a \cdot R_1). \tag{25}$$

Another criterion is the software reliability, which is defined as the probability that no failure occurs during time interval $(T,T+\Delta T]$ given that the software is released at time T. Considering that the reliability status of software generally does not change in operational phase, the reliability function is:

$$R_2(\Delta T|T) = \exp\left[-\lambda_d(T)\Delta T\right], \tag{26}$$

where $\lambda_d(T)$ is the instantaneous failure intensity at time T. For a given target R_2 for $R_2(\Delta T \mid T)$, the time for the software to reach R_2 can be solved as $\min_T\{T:R_2(\Delta T \mid T)\geq R_2\}$.

5.2. Software release policy based on cost criterion

For a basic FDP model with mean value function $m(t)$, the following cost model is frequently used [43]:

$$C(T) = c_1 m(T) + c_2\left(m(\infty) - m(T)\right) + c_3 T, \tag{27}$$

where c_1 is the expected cost of removing a fault during testing, c_2 is the expected cost of removing a fault in the field and c_3 is the expected cost per unit time of testing.In practice, the cost of removing a fault in field is generally greater than that during testing, thus we assume $c_2>c_1$.

When the correction process is incorporated, the following cost model can be constructed:

$$C(T) = c_1 m_r(T) + c_2\left(m_d(\infty) - m_r(T)\right) + c_3 T, \tag{28}$$

where $m_r(T)$ is the total number of corrected faults at the time of release T, and $m_d(\infty)-m_r(T)$ is the number of uncorrected faults that includes both the undetected faults $m_d(\infty)-m_d(T)$ and the detected-but-not-corrected faults $m_d(T)-m_r(T)$. By minimizing the cost model with respect to T, the optimal release time T_c under the proposed framework can be obtained.

Theorem 1: Under the proposed models in Section 3, the time T_c which minimizes $C(T)$ exists. Specifically, there exist $2k(k\geq 0)$ non-negative numbers $0 < z_1 \leq z_2 \leq \ldots \leq z_{2k} < +\infty$ which satisfy that $C(T)$ increases on $\left[z_{2j}, z_{2j+1}\right)$ and decreases on $\left[z_{2j+1}, z_{2j+2}\right)$ with $j=0,\ldots,k$, $z_0=0$ and $z_{2k+1}=+\infty$. The optimal T_c is determined as $T_c = \arg\min_{T\in\{0,z_2,\ldots,z_{2k}\}} C(T)$.

Proof: We just need to prove that there exists a T_s such that $C'(T)$ is positive for $T>T_s$. Since $C(T) = c_1 m_r(T) + c_2\left(m_d(\infty) - m_r(T)\right) + c_3 T$, we have:

$$C'(T) = c_3 - (c_2 - c_1)\lambda_r(T). \tag{29}$$

Clearly, $C'(0)=c_3>0$, indicating that $C(T)$ is increasing when T is close to zero. We shall prove that $\lambda_r(T)$ approaches 0 (or $C'(T)$ approaches c_3) when T approaches $+\infty$. If so, $C(T)$ is increasing when T is close to 0 or approaches $+\infty$. Consequently, if $C(T)$ has any stationary point, it must have even number of stationary points $0<z_1\leq z_2\leq\cdots\leq z_{2k}<+\infty$ such that $C(T)$ increases on on $\left[z_{2j}, z_{2j+1}\right)$ and decreases on $\left[z_{2j+1}, z_{2j+2}\right)$ for $j=0,\ldots,k$, $z_0=0$ and $z_{2k+1}=+\infty$. In the following, we shall show that $\lambda_r(T)$ approaches 0 when T approaches $+\infty$ under the three proposed models.

If the paired model under constant debugging lag assumption is used, from (16) we have:

$$\lambda_r(T) = pabe^{-b(T-\delta)} + ab(1-p)(bT - 2b\delta)e^{-b(T-2\delta)}, T \geq 2\delta. \tag{30}$$

When T approaches $+\infty$, $\lambda_r(T)$ approaches 0.

For the paired model with time-dependent debugging lags, according to (18) we have:

$$\lambda_d(T) = abe^{(-bT)} - a(1-p)(b + bcT - b^2T - (b^2cT^2)/2)e^{(-bT)}. \tag{31}$$

It can be seen that $\lambda_d(T)$ approaches 0 when T approaches $+\infty$. Moreover, we have:

$$\lambda_r(T) = \lambda_d\left(T - \frac{1}{b}\ln(1+\gamma T)\right)\left(1 - \frac{\gamma}{b(1+\gamma T)}\right).$$

As $T - b^{-1}\ln(1+\gamma T)$ approaches $+\infty$ when T approaches $+\infty$ for $b>c$, we can see that $\lambda_r(T)$ approaches 0 for $T\rightarrow+\infty$.

For the paired model under exponentially distributed random debugging lags, we have:

$$\lambda_r(T) = c\left(m_d(T) - m_r(T)\right). \tag{32}$$

Both $m_d(T)$ and $m_r(T)$ approach a as T approaches $+\infty$. Thus $\lambda_r(T)$approaches 0 when T approaches $+\infty$.

5.3. Software release policy based on mixed criterion

When both reliability requirements and the total cost are considered, we determine the optimal release time T^* that minimizes the total cost under the reliability constraint. Accordingly, the problem can be formulated as:

$$\text{Minimize } C(T) = c_1 m_r(T) + c_2\left(m_d(\infty) - m_r(T)\right) + c_3 T$$

$$\text{Subject to } R_1(T) = \frac{m_r(T)}{a} \geq R_1 \text{ (or } R_2(\Delta T|T) = \exp\left[-\lambda_d(T)\Delta T\right] \geq R_2).$$

When the reliability constraint $R_1(T)$ is used, we can divide the time axis $[0,\infty)$ into two types of intervals such that $C(T)$ increases on type 1 intervals and decreases on type 2 intervals. The candidates for T^* comprise of the minimum T on each type 1 interval that satisfies $R_1(T)\geq R_1$. Then, T^* is the one among all the candidates that leads to the lowest cost.

When the reliability constraint $R_2(\Delta T \mid T)$ is used, we can split the time axis $[0,\infty)$ into four types of intervals such that both $R_2(\Delta T \mid T)$ and $C(T)$ increase on type 1 intervals, both $R_2(\Delta T \mid T)$ and $C(T)$ decrease on type 2 intervals, $R_2(\Delta T \mid T)$ increases while $C(T)$ decreases on type 3 intervals, and $R_2(\Delta T \mid T)$ decreases while $C(T)$ increases on type 4 intervals. The candidates for T^* comprise of the minimum T in each type 1 interval that satisfies $R_2(\Delta T \mid T)\geq R_2$, the maximum T in each type 2 interval that satisfies $R_2(\Delta T \mid T)\geq R_2$, the end points of type 3 intervals which satisfy $R_2(\Delta T \mid T)\geq R_2$, and the initial points of type 4 intervals which satisfy $R_2(\Delta T \mid T)\geq R_2$. The optimal release time T^* is the one corresponding to the lowest cost.

5.4. Numerical examples

For illustration, we consider the paired FDP&FCP model with constant debugging lag that fits the dataset 1 in Section 4. The model parameters are a=199.27, b=0.00717, δ=24.78 and p=0.382. In addition, we assume c_1=\$300, c_2=\$2000, c_3=\$10, ΔT=12, R_1=0.95 and R_2=0.95. In the following, we present the optimal release time that minimizes the cost with specific reliability constraints.

1) Considering cost criterion and reliability target R_1.

From (28), the testing cost under our parameter settings is:

$$C(T) = 372500 - 1700 m_r(T) + 10T. \tag{33}$$

On the other hand, the correction process model with given parameters is:

$$m_r(T) = \begin{cases} 0, T < 24.78 \\ 76.12 - 90.92\mathrm{e}^{-0.00717T}, 24.78 \leq T < 49.56 \\ 199.27 - \left(204.18 + 1.26T\right)\mathrm{e}^{-0.00717T}, T \geq 49.56 \end{cases}.$$

By substituting $m_r(T)$ into (33), it can be derived that $C(T)$ increases on [0,24.78], decreases on (24.78,1030.45) and increases on $[1030.45,\infty)$. As can be verified, $R_1(0)$<0.95, $R_1(1030.45)$>0.95. According to the analysis in the preceding section, the optimal release time is $T_1^* = 1030.45$. Correspondingly, the optimal software testing costis $C\left(T_1^*\right) = 45624.87$.

2) Considering cost criterion and reliability target R_2.

When $R_2(\Delta T \mid T)$ is used as the reliability constraint, we can derive the following detection rate according to the specified model parameters:

$$\lambda_d(T) = \begin{cases} 0.5458\mathrm{e}^{-0.00717T}, T < 24.78 \\ (0.3584 + 0.0076T)\mathrm{e}^{-0.00717T}, T \geq 24.78 \end{cases}.$$

It can be verified that $\lambda_d(T)$ decreases on [0,24.78), increases on [24.78,92.07), and decreases on $[92.07,\infty)$. Accordingly, $R_2(\Delta T \mid T)$ increases on [0,24.78), decreases on [24.78,92.07), and increases on $[92.07,\infty)$. Referring to the analysis in Section 5.3, the axis $[0,\infty)$ can be divided into a type 1 interval [0,24.78), a type 2 interval [24.78,92.07), a type 3 interval [92.07,1030.45) and a type 1 interval $[1030.45,+\infty)$. Because $R_2(\Delta T \mid T)$<0.95 for all $T\in[0,92.07)$, there is no permissible T in this interval. Therefore, the candidates for the optimal T^* are the right endpoint 1030.45 of the type 3 interval and the minimum T in $[1030.45,+\infty)$ that satisfies $R_2(\Delta T \mid T)\geq 0.95$. Because $R_2(\Delta T \mid 1030.45)$=0.93<0.95, the optimal T_2^* is found as $\arg\min_T \left\{T \geq 1030.45, R(\Delta T|T) \geq 0.95\right\} = 1056.81$. The optimal software testing cost is $C\left(T_2^*\right) = 45645.38$, which is slightly larger than that in the last case. An illustration of the optimal release policies under two scenarios is given in Fig. 4.

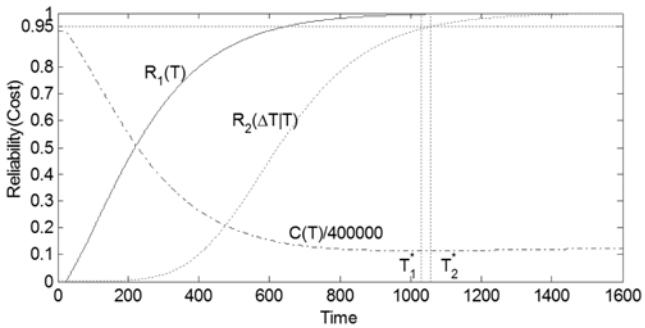


Fig. 4. Variation of normalized total cost function and software reliability functions with testing time

6. Conclusion

In this paper, we proposed a framework for the software reliability growth modeling. The software testing process was considered as a paired fault detection and correction process, and the faults during the testing were classified into leading and dependent faults according to their detectability. The leading faults can be detected and corrected directly, whereas the dependent faults can only be detected until the corresponding leading faults are corrected. For both types of faults, the FCP was modeled as a delayed FDP. In addition, the FDP of dependent faults depended on the FCP of leading faults. Special paired FDP&FCP models were derived under the proposed framework with different assumptions on the debugging lag. The application to two real software testing datasets revealed the effectiveness and the superiority of the proposed models over existing ones. Under this framework, the optimal software release policy was investigated considering cost and reliability requirements.

As a direction for future studies, the proposed modeling framework can be extended to incorporate other information or adapt to other testing environments. For instance, Bayesian technique can be used to incorporate prior information and update model parameters when more information is available. In addition, the imperfect fault correction or the fault introduction phenomenon can be incorporated, as it is common for debuggers to make mistakes with fault correction.

Acknowledgements

The research is supported by the NSFC under grant number 71671016 and 71231001 and 71420107023,and by the Fundamental Research Funds for the Central Universities of China FRF-BR-15-001B.

References

1. Boland P J and Chuív N N. Optimal times for software release when repair is imperfect. Statistics & Probability Letters 2007; 77(12): 1176-1184, https://doi.org/10.1016/j.spl.2007.03.004.
2. Chang Y-C and Liu C-T. A generalized JM model with applications to imperfect debugging in software reliability. Applied Mathematical Modelling 2009; 33(9): 3578-3588, https://doi.org/10.1016/j.apm.2008.11.018.
3. Febrero F, Calero C, and Ángeles Moraga M. Software reliability modeling based on ISO/IEC SQuaRE. Information and Software Technology 2016; 70: 18-29, https://doi.org/10.1016/j.infsof.2015.09.006.
4. Goel A L and Okumoto K. Time-dependent error-detection rate model for software reliability and other performance measures. IEEE Transactions on Reliability 1979; 28(3): 206-211, https://doi.org/10.1109/TR.1979.5220566.
5. Gokhale S S, Lyu M R, and Trivedi K S. Analysis of software fault removal policies using a non-homogeneous continuous time Markov chain. Software Quality Journal 2004;12(3): 211-230, https://doi.org/10.1023/B:SQJO.0000034709.63615.8b.
6. Goseva-Popstojanova K and Trivedi K S. Failure correlation in software reliability models. IEEE Transactions on Reliability 2000; 49(1): 37-48, https://doi.org/10.1109/24.855535.
7. Gutjahr W J. A reliability model for nonhomogeneous redundant software versions with correlated failures. Computer Systems Science and Engineering 2001;16(6): 361-370.
8. Hu Q, Xie M, Ng S H, and Levitin G. Robust recurrent neural network modeling for software fault detection and correction prediction. Reliability Engineering & System Safety 2007;92(3): 332-340. https://doi.org/10.1016/j.ress.2006.04.007
9. Huang C-Y and Huang W-C, Software reliability analysis and measurement using finite and infinite server queueing models. IEEE Transactions on Reliability 2008; 57(1): 192-203, https://doi.org/10.1109/TR.2007.909777.
10. Huang C-Y, Kuo S-Y, and Lyu M R. An assessment of testing-effort dependent software reliability growth models. IEEE Transactions on Reliability 2007; 56(2): 198-211, https://doi.org/10.1109/TR.2007.895301.
11. Huang C-Y and Lin C-T. Software reliability analysis by considering fault dependency and debugging time lag. IEEE Transactions on Reliability 2006;55(3): 436-450. https://doi.org/10.1109/TR.2006.879607
12. Huang C-Y and Lyu M R, Optimal release time for software systems considering cost, testing-effort, and test efficiency. IEEE Transactions on Reliability 2005; 54(4): 583-591, https://doi.org/10.1109/TR.2005.859230.
13. Inoue S and Yamada S. Generalized discrete software reliability modeling with effect of program size. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 2007; 37(2): 170-179, https://doi.org/10.1109/TSMCA.2006.889475.
14. Jain M and Priya K. Software reliability issues under operational and testing constraints. Asia-Pacific Journal of Operational Research 2005; 22(01): 33-49, https://doi.org/10.1142/S021759590500042X.
15. Jha P, Gupta D, Yang B, and Kapur P. Optimal testing resource allocation during module testing considering cost, testing effort and reliability. Computers & Industrial Engineering 2009; 57(3): 1122-1130, https://doi.org/10.1016/j.cie.2009.05.001.
16. Jia L, Yang B, Guo S, and Park D H. Software reliability modeling considering fault correction process. IEICE Transactions on Information and Systems 2010; 93(1): 185-188, https://doi.org/10.1587/transinf.E93.D.185.
17. Kapur P, Goswami D, Bardhan A, and Singh O. Flexible software reliability growth model with testing effort dependent learning process. Applied Mathematical Modelling 2008; 32(7): 1298-1307, https://doi.org/10.1016/j.apm.2007.04.002.
18. Kapur P and Younes S. Software reliability growth model with error dependency. Microelectronics Reliability 1995; 35(2): 273-278, https://doi.org/10.1016/0026-2714(94)00054-R.
19. Kim H S, Park D H, and Yamada S. Bayesian optimal release time based on inflection S-shaped software reliability growth model. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences 2009; 92(6): 1485-1493, https://doi.org/10.1587/transfun.E92.A.1485.
20. Lee C H, Kim Y T, and Park D H. S-shaped software reliability growth models derived from stochastic differential equations. IIE transactions 2004; 36(12): 1193-1199, https://doi.org/10.1080/07408170490507792.
21. Li X, Xie M, and Ng S H. Sensitivity analysis of release time of software reliability models incorporating testing effort with multiple change-points. Applied Mathematical Modelling 2010; 34(11): 3560-3570, https://doi.org/10.1016/j.apm.2010.03.006.
22. Lin C-T and Huang C-Y., Enhancing and measuring the predictive capabilities of testing-effort dependent software reliability models. Journal of Systems and Software 2008; 81(6): 1025-1038, https://doi.org/10.1016/j.jss.2007.10.002.
23. Lin C-T and Huang C-Y. Staffing level and cost analyses for software debugging activities through rate-based simulation approaches. IEEE Transactions on Reliability 2009; 58(4): 711-724, https://doi.org/10.1109/TR.2009.2019669.
24. Lin C-T and Li Y-F. Rate-based queueing simulation model of open source software debugging activities. IEEE Transactions on Software Engineering 2014; 40(11): 1075-1099, https://doi.org/10.1109/TSE.2014.2354032.
25. Lo J-H and Huang C-Y. An integration of fault detection and correction processes in software reliability analysis. Journal of Systems and Software 2006; 79(9): 1312-1323, https://doi.org/10.1016/j.jss.2005.12.006.
26. Lyu M R. Handbook of Software Reliability Engineering. New York: McGraw-Hill, Inc., 1996.
27. Musa J D, Iannino A, and Okumono K. Software Reliability, Measurement, Prediction and Application. New York: McGraw-Hill, Inc, 1987.
28. Okamura H and Dohi T. Unification of software reliability models using Markovian arrival processes, in Proceedings of 17th Pacific Rim International Symposium on Dependable Computing (PRDC) 2011: 20-27, https://doi.org/10.1109/prdc.2011.12.

29. Okamura H, Dohi T, and Osaki S. Software reliability growth models with normal failure time distributions. Reliability Engineering & System Safety 2013; 116: 135-141, https://doi.org/10.1016/j.ress.2012.02.002.
30. Peng R, Li Y-F, Zhang J-G, and Li X. A risk-reduction approach for optimal software release time determination with the delay incurred cost. International Journal of Systems Science 2015; 46(9): 1628-1637, https://doi.org/10.1080/00207721.2013.827261.
31. Peng R, Li Y-F, Zhang W, and Hu Q. Testing effort dependent software reliability model for imperfect debugging process considering both detection and correction. Reliability Engineering & System Safety 2014;126: 37-43, https://doi.org/10.1016/j.ress.2014.01.004.
32. Pietrantuono R, Russo S, and Trivedi K S. Software reliability and testing time allocation: An architecture-based approach. IEEE Transactions on Software Engineering 2010; 36(3): 323-337, https://doi.org/10.1109/TSE.2010.6.
33. Rana R, Staron M, Berger C, Hansson J, Nilsson M, Törner F, et al., Selecting software reliability growth models and improving their predictive accuracy using historical projects data. Journal of Systems and Software 2014; 98: 59-78, https://doi.org/10.1016/j.jss.2014.08.033.
34. Schneidewind N F. Analysis of error processes in computer software, in Proceedings of 1975 International Conference on Reliable Software 1975: 337-346, https://doi.org/10.1145/800027.808456.
35. Shatnawi O. Discrere time modelling in software reliability engineering-A unified approach. Computer Systems Science and Engineering 2009;24(6): 391.
36. Shatnawi O. Measuring commercial software operational reliability: an interdisciplinary modelling approach. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014;16(4): 585-594.
37. Shibata K, Rinsaka K, Dohi T, and Okamura H. Quantifying software maintainability based on a fault-detection/correction model, in Proceedings of 13th Pacific Rim International Symposium on Dependable Computing (PRDC2007) 2007: 35-42, https://doi.org/10.1109/PRDC.2007.46.
38. Shinohara Y, Nishio Y, Dohi T, and Osaki S. An optimal software release problem under cost rate criterion: artificial neural network approach. Journal of Quality in Maintenance Engineering 1998; 4(4): 236-247, https://doi.org/10.1108/13552519810233967.
39. Tamura Y and Yamada S. A flexible stochastic differential equation model in distributed development environment. European Journal of Operational Research 2006; 168(1): 143-152, https://doi.org/10.1016/j.ejor.2004.04.034.
40. Ullah N, Morisio M, and Vetro A. A comparative analysis of software reliability growth models using defects data of closed and open source software, in Proceedings of 35th Annual IEEE Software Engineering Workshop 2012: 187-192, https://doi.org/10.1109/sew.2012.26.
41. Wang L, Hu Q, and Liu J. Software reliability growth modeling and analysis with dual fault detection and correction processes. IIE Transactions 2016; 48(4): 359-370, https://doi.org/10.1080/0740817X.2015.1096432.
42. Wu Y, Hu Q, Xie M, and Ng S H. Modeling and analysis of software fault detection and correction process by considering time dependency. IEEE Transactions on Reliability 2007; 56(4): 629-642, https://doi.org/10.1109/TR.2007.909760.
43. Xie M. Software reliability modelling. Singapore: World Scientific, 1991, https://doi.org/10.1142/1390.
44. Xie M, Hu Q, Wu Y, and Ng S H. A study of the modeling and analysis of software fault-detection and fault-correction processes. Quality and Reliability Engineering International 2007; 23(4): 459-470, https://doi.org/10.1002/qre.827.
45. Yamada S, Ohba M, and Osaki S. S-shaped software reliability growth models and their applications. IEEE Transactions on Reliability 1984; 33(4): 289-292, https://doi.org/10.1109/TR.1984.5221826.
46. Yang B, Guo S, Ning N, and Huang H-Z. Parameter estimation for software reliability models considering failure correlation, in Proceedings of Annual Reliability and Maintainability Symposium (RAMS 2008) 2008: 405-410, https://doi.org/10.1109/rams.2008.4925830.
47. Zhang X and Pham H. Comparisons of nonhomogeneous Poisson process software reliability models and its applications. International Journal of Systems Science 2000; 31(9): 1115-1123, https://doi.org/10.1080/002077200418397.
48. Zhang X and Pham H. Predicting operational software availability and its applications to telecommunication systems. International Journal of Systems Science 2002; 33(11): 923-930, https://doi.org/10.1080/0020772021000023022.
49. Zhang X, Teng X, and Pham H. Considering fault removal efficiency in software reliability assessment. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 2003; 33(1): 114-120, https://doi.org/10.1109/TSMCA.2003.812597.
50. Zhao J, Liu H-W, Cui G, and Yang X-Z. Software reliability growth model with change-point and environmental function. Journal of Systems and Software 2006; 79(11): 1578-1587, https://doi.org/10.1016/j.jss.2006.02.030.

Rui PENG
Donlinks School of Economics & Management
University of Science & Technology Beijing, China

Qingqing ZHAI
Department of Industrial and Systems Engineering
National University of Singapore, Singapore

E-mails: pengrui1988@ustb.edu.cn, isezq@nus.edu.sg

Article citation info:
SKRICKIJ V, BOGDEVIČIUS M, ZYGIENĖ R. Evaluation of the spur gear condition using extended frequency range. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 476–484, http://dx.doi.org/10.17531/ein.2017.3.19.

Viktor SKRICKIJ
Marijonas BOGDEVIČIUS
Rasa ŽYGIENĖ

EVALUATION OF THE SPUR GEAR CONDITION USING EXTENDED FREQUENCY RANGE

OCENA STANU PRZEKŁADNI ZĘBATEJ Z WYKORZYSTANIEM ROZSZERZONEGO ZAKRESU CZĘSTOTLIWOŚCI

The paper focuses on working out an algorithm for spur gear condition monitoring, based on the results of numerical simulation. The nonlinear mathematical model has been used for investigation of the dynamic parameters of the cylindrical spur gear with defective teeth. Backlash between gear teeth, backlash in bearings, time-varying mesh stiffness, and variations of the centre distance have been evaluated in the model. Diagnostic parameters suitable for determining the condition of the gears under investigation have been established. Frequency intervals mostly affected by changes in diagnostic parameters under damage have been found. An algorithm for diagnostics based on mathematical modelling, vibro-acoustic, and acoustic emission methods, and wavelet transform has been worked out.

Keywords: *mathematical model; spur gear; acoustic emission; vibro-acoustic signal, wavelet transform, condition monitoring.*

Celem artykułu było opracowanie algorytmu monitorowania stanu przekładni zębatej w oparciu o wyniki symulacji numerycznej. Przedstawiono nieliniowy model matematyczny, który wykorzystano do badania parametrów dynamicznych przekładni zębatej walcowej z uszkodzonymi zębami. Za pomocą przedstawionego modelu oceniano luz pomiędzy zębami przekładni, luz w łożyskach, zmienną w czasie sztywność zazębienia oraz zmiany odległości osi. Ustalono parametry diagnostyczne odpowiednie dla określenia stanu technicznego badanych przekładni. Znaleziono przedziały częstotliwości odpowiadające zmianom parametrów diagnostycznych wynikającymi z uszkodzenia. Opracowano algorytm diagnostyczny oparty na modelowaniu matematycznym, metodach emisji wibroakustycznej i emisji akustycznej oraz transformacie falkowej.

Słowa kluczowe: *model matematyczny; przekładnia zębata; emisja akustyczna; sygnał wibroakustyczny, transformata falkowa, monitorowanie stanu.*

1. Introduction

Improving equipment performance and increasing productivity are the major tasks for the operation of modern multi-purpose technological equipment. Spur gears are widely used in such equipment and gear damage may lead the whole system to failure, so condition monitoring is a commonly used way to improve reliability of equipment usage. Experimental research is the method most frequently used for the investigation of dynamic parameters of equipment. However, an experimental way cannot be the only technique for investigating off all possible combinations of operating parameters such as speeds and loads due to high costs and time spent for determining the causes. Mathematical modelling of a physical system employing numerical methods has become an effective alternative.

Spur gear mathematical models can be divided into analytical and numerical. Review of analytical models, where linear systems have been investigated, are presented in Ozguven and Houser [23]. Nowadays, usage of these models is limited, they are often used for calculations of the eigenvalues, also they can be used in systems where main research objects are not spur gears. One of the first nonlinear model was presented by Utagawa [32]. Author found out dynamic loads in the gear, by using time-variant stiffness. All data was compared with experimental investigation. Kahraman and Singh [9] investigated gear dynamics, when there is a backlash between gear teeth. Maliha et al. [18] investigated multibody system: gear, shafts, bearings, also backlash was included, others similar models were reviewed in [35]. Kiekbusch and Howard [11] presented equations for rotational gear mesh stiffness calculation in their investigation. Saxena et al. [27] used modified gear mesh stiffness model, which was based on potential energy method proposed by Yang and Lin [37]. The energy stored in meshing gear system was assumed to include four components: Hertzian energy, bending energy, axial compressive energy and shear energy. Multi-level gear models are presented in [2, 4, 35]. Cheon [2] has used Fourier series for mesh stiffness evaluation, backlash between gear teeth was taken into account and damping was included. Fakhfakh et al. [4] stiffness evaluated using time-variant function. Vaishya and Singh [33] proposed model where sliding friction in mesh was included. He et al. [7] compared five different friction models, authors investigated friction influence on gear dynamics. Amabili and Rivola [1] included mesh damping in their work. Kuang and Lin [12], Wojnarowski and Onishchenko [36] investigated gear teeth wear problem. Walha et al. [34] took bearing deformations and backlashes into account, meshing stiffness was proposed as a function. Frolov and Kosarev [5] investigated which parameters have most significant effect to gear dynamics. It was found that main parameters in gear dynamics are mesh stiffness, pitch errors, teeth surface errors. Jia and Howard [8], Fakher et al. [3] investigated pitting problem. Litak and Friswell [15] investigated teeth breakage, pitch errors. Zouari et al. [40] investigated teeth crack influence on mesh stiffness, authors changed crack level, crack direction, FEM was used for modelling.

Mohammeda et al. [22] investigated tree different types of teeth cracks. First, crack is over all teeth and has constant value on every teeth width, second, crack depth is crack depth is distributed according to a parabolic function; third, crack is growing in all directions. As diagnostic parameters RMS and Excess were used, mesh stiffness was calculated using potential energy. Ma et al. [17] investigated a phenomenon of cracked teeth when the incoming tooth pair enters contact earlier than the theoretical start of contact, and the outgoing tooth pair leaves contact later than the theoretical end of contact. It was found that for simulated signals, sideband frequencies, statistical features and instantaneous energy can all reflect the fault features of the cracked gear.

Various methods, including the measurement of the vibro-acoustic signal (VS), oil debris method, sound measurement, acoustic emission (AE) and temperature measurement could be used for experimental identification of gear condition. VS and AE methods enable to identify variations in the condition and the type of a defect. The VS is a well-known method applied for testing the rotor system and widely described by Taylor and Kirkland [31]. This method has been extensively employed in the diagnostics of the rotor system. AE was originally developed for non-destructive testing of static structures; however, over the years, its application has been extended to health monitoring of rotating machines and bearings [20].

All measured data processing methods could be divided into three main domains [30]: time domain, frequency domain and time-frequency domain. Diagnostic parameters used in time domain are presented in Fig. 1, more information about these parameters can be found in [6, 10, 14, 16, 19, 21, 24, 25, 26, 38, 39]. Some of them can be also used in frequency and in time-frequency domains. Time synchronous averaging can be applied and filtering can be used, if signal has a lot of noise. Difference and residual signals are needed for some diagnostic parameters used in gear diagnostics [35].

In frequency domain, we can establish the type of the defect. Very powerful tool is signal analysis in time-frequency domain, it enables investigation of a signal not only according to frequencies, but on the time scale as well. It could be useful when a mechanism works at non-stationary modes. Methods generally used to process a signal in time-frequency scales are Short Time Fourier Transform (STFT), Wigner-Ville distribution, Choi-Williams distribution and Wavelet transform. Wavelet is not a direct time-frequency representation, but time-scale. If compared to STFT, Wavelet uses narrow time windows at high frequencies and wide time windows at low frequencies, using Wavelet transform computing time is decreasing. Skrickij et al. [28] showed that using wavelet transform and an extended frequency range, AE and VS signal monitoring of the gear unit is much more sensitive, and the occurrence of teeth faults and their growth can be recognized at an earlier stage.

The conducted analysis of mathematical models for spur gears has demonstrated that nonlinear dynamic systems must be examined in order to make the models for investigation of the dynamic parameters of defective gears. Also, the models of the gears must have the assessed backlashes in the bearings, and between gears. Mesh stiffness should be simulated as a time function and depend on the number of the pairs of gear teeth involved in the mesh as well as on the place of contact. Researchers do not focus on the defective top-part of the gear teeth, they mostly focus on teeth crack and pitting problem. The defect in the top-part of teeth appears in gearboxes when gears are moving in respect to each other during gear shifting.

In this paper diagnostic parameters suitable for determining the condition of the gears under investigation are established, frequency intervals with the most significant increase in the diagnostic parameter values are determined. Also it is found that diagnostic parameters are most sensitive to the occurrence and expansion of the defect when using the proposed mathematical model.

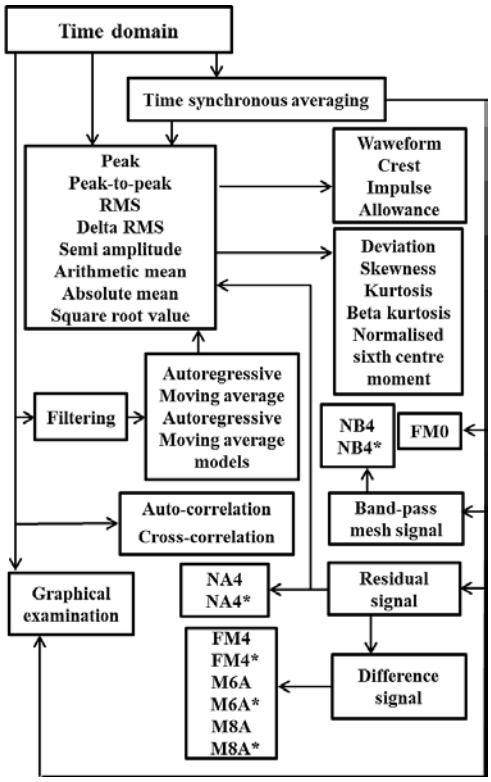


Fig. 1. Data processing methods

2. Mathematical model for defective spur gear

To investigate spur gear defect influence on its dynamical characteristics, nonlinear mathematical model was used. Gear is modelled as 2D system, only rotation around x axis and displacements in y, z directions (Fig. 2) are taken into account; lubrication has not been included in the model; gear teeth deformations are taken into account; Kuang and Yang method [13] has been applied for time-variant mesh stiffness evaluation; shafts rotation deformations have been taken into account, bending and contact deformations have not been included; contact deformations in bearings have been estimated; backlash in bearings between rotation elements and inner and outer races has been estimated; backlash between gear teeth has been estimated; centre distance error has been evaluated; variations in centre distance due to gear work has been evaluated.

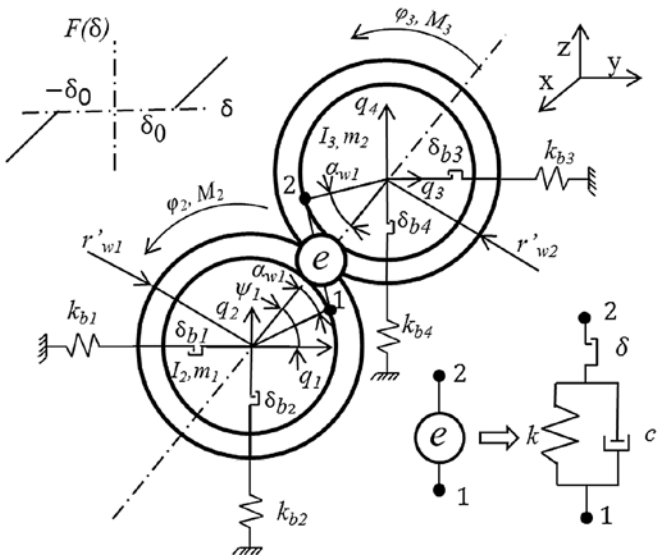


Fig. 2. Dynamic model of gear with backlash

Errors may occur in the process of manufacturing, and number of faults is growing during gear operation, backlashes are getting wider and the flexibility of units is rising, centre distance is changing. The centres of gears are moving because rotation torques are applied. As for the operating mode of the mechanism, bearing flexibility is growing thus causing further changes in centre distance. Thus, a mathematical model of the gear train where variations in centre distance are evaluated was presented in [29].

Dynamic model of spur gear is presented in Fig. 2. The displacements of points 1 and 2 along a straight line, which is a tangent line to the circles of both gears (Fig. 2), are calculated as follows:

$$u_1 = r'_{w1} \cdot \varphi_2 - q_1 \cdot \sin(\psi_1 - \alpha_{w1}) + q_2 \cdot \cos(\psi_1 - \alpha_{w1}) \quad (1)$$

$$u_2 = -r'_{w2} \cdot \varphi_3 - q_3 \cdot \sin(\psi_1 - \alpha_{w1}) + q_4 \cdot \cos(\psi_1 - \alpha_{w1}) \quad (2)$$

$$\delta = u_2 - u_1 \quad (3)$$

where r'_{w1} is radii of the pitch circles; φ_2, φ_3 are rotation angles; ψ_1 is orientation angle; α_{w1} is pressure angle; q_i is displacement of gear, for i = 1, 2, 3, 4; parameters $\psi_1, \alpha_{w1}, r'_{wi}$ in this case are not constants, evaluation of these parameters are presented in [29]. The velocities of points 1 and 2 are calculated as follows:

$$\dot{u}_1 = r'_{w1} \cdot \dot{\varphi}_2 - \dot{q}_1 \cdot \sin(\psi_1 - \alpha_{w1}) + \dot{q}_2 \cdot \cos(\psi_1 - \alpha_{w1}) \quad (4)$$

$$\dot{u}_2 = -r'_{w2} \cdot \dot{\varphi}_3 - \dot{q}_3 \cdot \sin(\psi_1 - \alpha_{w1}) + \dot{q}_4 \cdot \cos(\psi_1 - \alpha_{w1}) \quad (5)$$

$$\dot{\delta} = \dot{u}_2 - \dot{u}_1 \quad (6)$$

where $\dot{\varphi}_i = \frac{d\varphi_i}{dt}$ is angular velocity, for i = 1, 2, 3, 4; $\dot{q}_i = \frac{dq_i}{dt}$ is linear velocity. The influence of the backlash is showed in Fig. 2, parameter δ is a backlash. The force acting on mesh is obtained from the expression:

$$F = -\mathrm{k}\tilde{\delta} - \mathrm{c} \cdot \dot{\delta} \quad (7)$$

where c is damping coefficient (Fig. 2); k is mesh stiffness determined by the method offered by Kuang and Yang [13], where $\tilde{\delta}$ denotes the excess of displacement in view of the backlash of gears:

$$\tilde{\delta} = \begin{cases} \delta - \delta_0, \ when\ \delta > \delta_0 \\ 0, when - \delta_0 \le \delta \le \delta_0 \\ \delta + \delta_0, when\ \delta < -\delta_0 \end{cases} \quad (8)$$

Stiffness of one tooth is evaluated using formulas:

$$K_i(r) = (A_0 + A_1 X_i) + (A_2 + A_3 X_i) \frac{r - r'_{wi}}{(1 + X_i)m} \quad (9)$$

where:

$$A_0 = 3.867 + 1.612 \cdot z_i - 0.02916 \cdot z_i^2 + 0.0001553 \cdot z_i^3 \quad (10)$$

$$A_1 = 17.060 + 0.7289 \cdot z_i - 0.01728 \cdot z_i^2 + 0.0000999 \cdot z_i^3 \quad (11)$$

$$A_2 = 2.637 - 1.222 \cdot z_i - 0.02217 \cdot z_i^2 + 0.0001179 \cdot z_i^3 \quad (12)$$

$$A_3 = -6.330 - 1.033 \cdot z_i + 0.02068 \cdot z_i^2 - 0.0001130 \cdot z_i^3 \quad (13)$$

where r is gear radii where load is applied; X_i is a parameter in this mathematical model, it is assumed, that X_i= 0; m is gear module; z_i is number of teeth in gear.

One tooth stiffness:

$$k_i = K_i(r) \cdot b \cdot 10^9 \quad (14)$$

where b is tooth width, if two teeth are in contact, mesh stiffness can be calculated:

$$k = \frac{k_1 \cdot k_2}{k_1 + k_2} \quad (15)$$

If four teeth are in contact, stiffness can be evaluated:

$$k = \frac{k_1 \cdot k_2}{k_1 + k_2} + \frac{k_3 \cdot k_4}{k_3 + k_4} \quad (16)$$

If gear tooth has top-part defect, time is decreasing, while four teeth are in contact. If tooth is cracked, stiffness of this tooth k_i is decreasing, meshing stiffness is decreasing too. If there is pitting problem stiffness of tooth k_i is decreasing, also a loss of contact is possible. It is very difficult to find the top part defect, because stiffness of tooth is not changing significantly (Fig. 4).

Bearing force is found to be:

$$F_{bi} = -k_b \cdot \tilde{q}_i^{1.5} \left(1 + a_b \cdot (1 - e_n^2) \frac{\dot{q}_i}{\dot{\Delta}}\right) \quad (17)$$

where k_b is bearing contact stiffness, $\dot{\Delta}$ is the penetration rate; a_b is the coefficient; e_n is the restitution coefficient, $\tilde{q}_i$ denotes the excess of displacement in view of the backlash of bearing:

$$\tilde{q}_i = \begin{cases} q_i - \delta_b, when\ q_i > \delta_b \\ 0, when - \delta_b \le q_i \le \delta_b \\ q_i + \delta_b, when\ q_i < -\delta_b \end{cases} \quad (18)$$

where δ_b is backlash in bearing.

Axial displacements of gears are determined as:

$$m_1 \cdot \ddot{q}_1 = \mathrm{F} \cdot \sin(\psi_1 - \alpha_{w1}) + F_{b1} \quad (19)$$

$$m_1 \cdot \ddot{q}_2 = -\mathrm{F} \cdot \cos(\psi_1 - \alpha_{w1}) + F_{b2} - m_1 \cdot g \quad (20)$$

$$m_2 \cdot \ddot{q}_3 = -\mathrm{F} \cdot \sin(\psi_1 - \alpha_{w1}) + F_{b3} \quad (21)$$

$$m_2 \cdot \ddot{q}_4 = \mathrm{F} \cdot \cos(\psi_1 - \alpha_{w1}) + F_{b4} - m_2 \cdot g \quad (22)$$

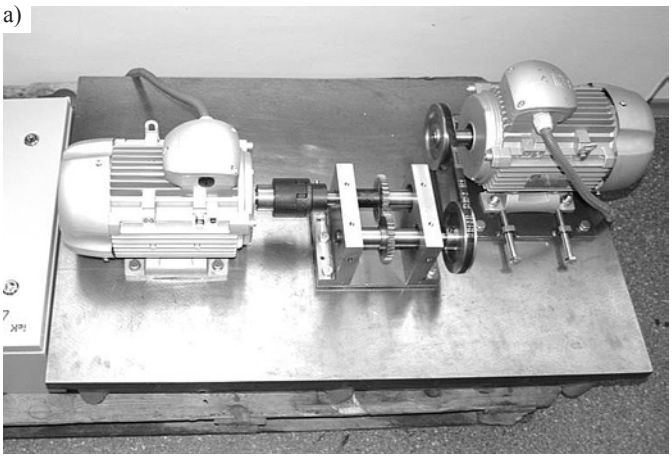
where b is gravity acceleration, m_1 is mass of gear, m_2 is mass of pinion.

Equation for gear rotation are presented in Skrickij and Bogdevicius [29]. Using the presented model different gear defects can be explored, such as teeth crack, pitting problem and teeth top defect.

3. Test Rig description and experimental procedure

This investigation involves experimental testing of spur gear to obtain its dynamic characteristics under different conditions. For this purpose, a series of test were carried out using test rig presented in Fig. 3, a). One of the engines was operating under motor mode while the other had to carry load. Sensors were placed on the bearing housing in the vertical direction and vibro-acoustic and acoustic emission signals were measured. To measure VS acquisition system MTX 1054, with AS-065 sensor, 100 mV/g (1 ... 15 000 Hz, ± 3 dB) were used, to measure AE signal the data acquisition system Mistras Pocket AE-2, with $R15\alpha$ sensor with resonant frequency of 75 kHz was applied. A cylindrical spur gear without lubrication was under investigation. Gear ratio – 1; number of teeth in gears z=30 gear module m=2.5 mm gear width b=10 mm. Revolutions are 1487 RPM and constant load.

a)



b)

Fig. 3. a) Test rig for spur gear dynamic research; b) gear with a defect

To reduce measuring error, each measurement was repeated five times. The examined cases cover a working gear and the defective gear with 1 and 2 mm removed from the tooth of the drive gear (Fig 3, b).

4. Results of mathematical simulation and model verification

In this part of paper simulation results as well as results of experimental investigation are presented and verification is made.

Modelling, considering three cases, including a working gear, a decrease in the height of the tooth of the gear in 1 mm and in 2 mm, is proposed. Detection of such defects is a complex problem and occurs only when the upper tooth leaves the mesh Fig.4.

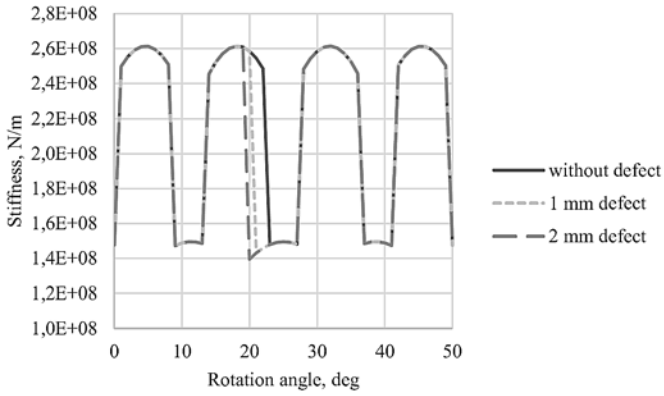


Fig. 4. The influence of a defect on meshing frequency

Discrete Wavelet Transform (DWT) and Daubechies 5 (DB5) Wavelet were selected to conduct mathematical modelling results and experiment. The presented Wavelet has good resolution in the domain of low frequencies. The AE signal is resolved into 9 intervals from 0 to 50 kHz (Fig. 5, b), the VS is resolved into 6 intervals from 0 to 5 kHz (Fig. 5, a). For investigation, the following diagnostic parameters were employed: X1 – Peak, X2 – Peak to Peak, X3 – RMS. Usage of these parameters gives good diagnostic results and minimal number of faulty diagnosis [28].

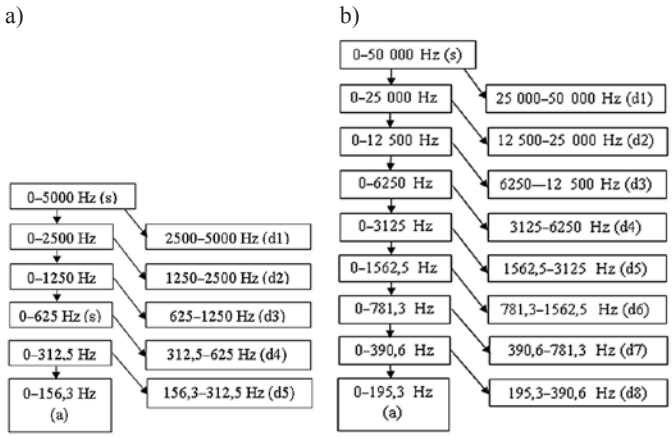


Fig. 5. The decomposition of the signal into frequency intervals: a) the Vibro-acoustic signal is resolved into intervals from 0 to 5 kHz; b) the Acoustic Emission signal is resolved into 9 intervals from 0 to 50 kHz

Using results of the mathematical simulation the values of diagnostic parameters (X1, X2, X3) are calculated in every frequency interval. The examination of the obtained results suggests that mostly the values of diagnostic parameters vary in frequency intervals d1 (2500-5000Hz) and d2 (1250-2500Hz). The X1 parameter has increased by 1.25 times from 2500 to 5000 Hz (d1) after introducing a defect of 1 mm, 1.43 times after introducing a defect of 2 mm. Same parameter has increased by 1.51 times after introducing a defect of 1 mm and 1.85 times after introducing a defect of 2 mm in frequency interval from 1250 to 2500 Hz (d2) (Fig. 6 a).

The X2 parameter has increased by 1.14 times (Fig. 6 b) from 2500 to 5000 Hz (d1) after introducing a defect of 1 mm, 1.28 times after introducing a defect of 2 mm. Same parameter has increased by 1.45 times introducing a defect of 1 mm and 1.80 times after introducing a defect of 2 mm in frequency interval from 1250 to 2500 Hz (d2). The increase of X3 parameter (Fig. 7) from 1250 to 5000 Hz wasn't significant.

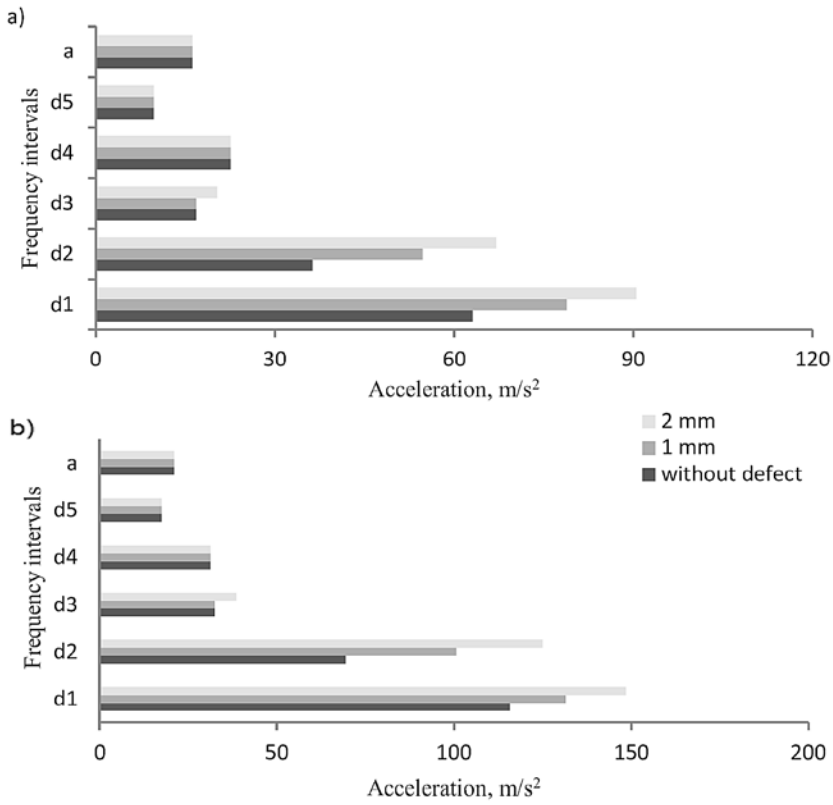


Fig. 6. Results of mathematical modelling, diagnostic parameter: a) X1 (Peak) b) X2 (Peak to Peak)

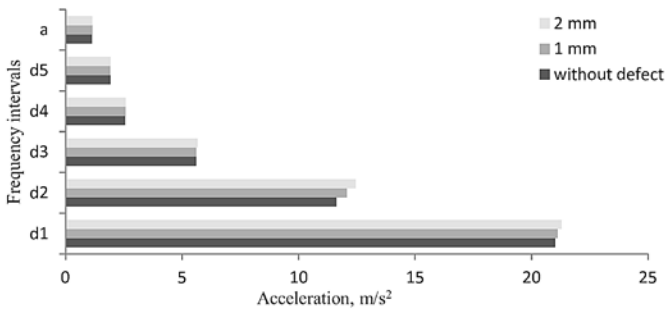


Fig. 7. Results of mathematical modelling, diagnostic parameter X3 (RMS)

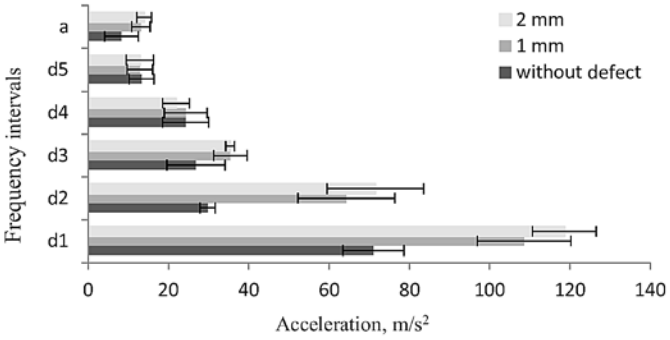


Fig. 8. The Vibro-acoustic signal of the gear, diagnostic parameter X1 (Peak)

Drive vibrations at three different conditions have been measured. The received data was processed by decomposing the signal into frequency intervals. VS signal was divided into 6 intervals (Fig. 5 a), AE signal – into 9 (Fig. 5 b). On the basis of the results of mathematical modelling, two basic diagnostic parameters X1, X2 have been chosen, and also, X3 was checked.

The use of RMS has not been approved. The results and research on mathematical modelling applying the AE method demonstrated a fault of a measuring device, and that 10-bit resolution was not enough.

Defects are best determined under high frequencies. Variations in parameter X1 are observed within the intervals of 1250-2500Hz and 2500- 5000Hz (Fig. 8). The X1 parameter has increased by 2.15 times from 1250 to 2500 Hz (d1) after introducing a defect of 1 mm, 2.40 times after introducing a defect of 2 mm. Same parameter has increased by 1.53 times in frequency interval from 2500 to 5000 Hz (d2) after introducing a defect of 1 mm and 1.67 times after introducing a defect of 2 mm (Fig. 8).

The X2 parameter has increased by 2.19 times from 1250 to 2500 Hz (d1) after introducing a defect of 1 mm, 2.38 times after introducing a defect of 2 mm. Same parameter has increased by 1.56 times from 2500 to 5000 Hz (d2) after introducing a defect of 1 mm and 1.71 times after introducing a defect of 2 mm (Fig. 9).

In parallel with measuring of the VS, AS was assessed. The obtained results are presented below. For interpretation of AE data in different frequency intervals, the same diagnostic parameters, as in the case of VS, are invoked.

The X1 parameter (Fig. 11) has increased by 1.20 times in d4 (3125 – 6250 Hz) frequency range, after introducing a defect of 1 mm, by 1.11 times in d5 (1562.5 – 3125 Hz) frequency range, by 1.08 times in

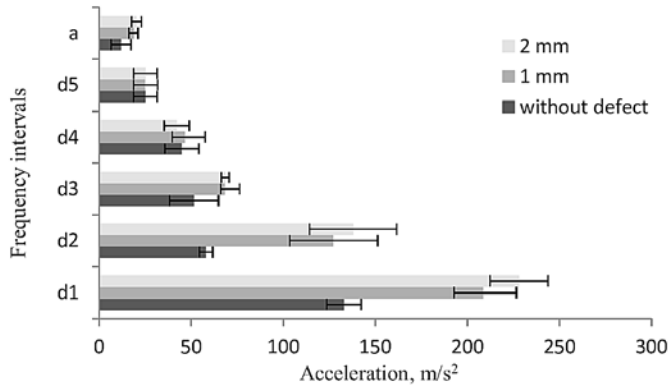


Fig. 9. The Vibro-acoustic signal of the gear, diagnostic parameter X2 (Peak to Peak)

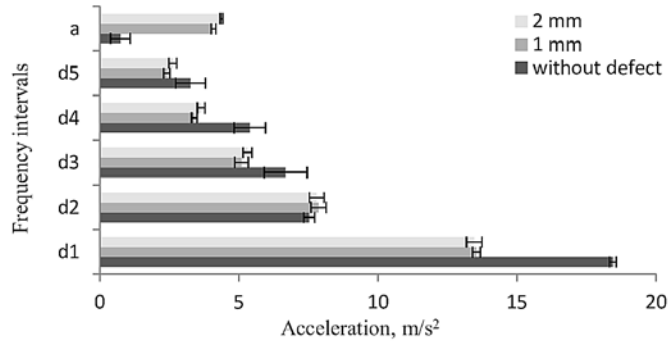


Fig. 10. The Vibro-acoustic signal of the gear, diagnostic parameter X3 (RMS)

d6 (781.3 – 1562.5 Hz) frequency range, by 3.88 times in d2 (12.5 – 25 kHz) frequency range, by 2.64 times in d1 (25 – 50 kHz) frequency range. After introducing a defect of 2 mm, X1 parameter has increased

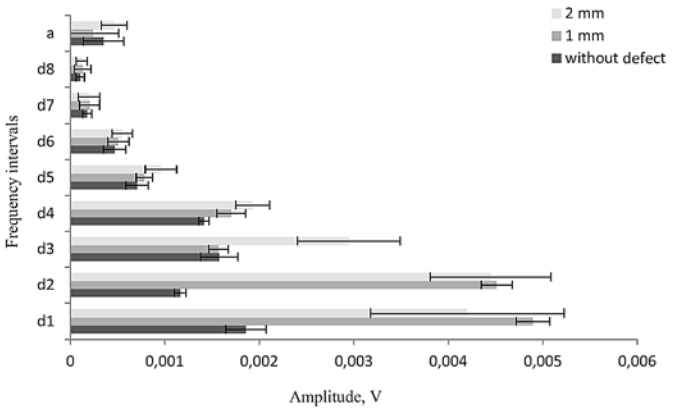


Fig. 11. The acoustic emission signal of the gear, diagnostic parameter X1 (Peak)

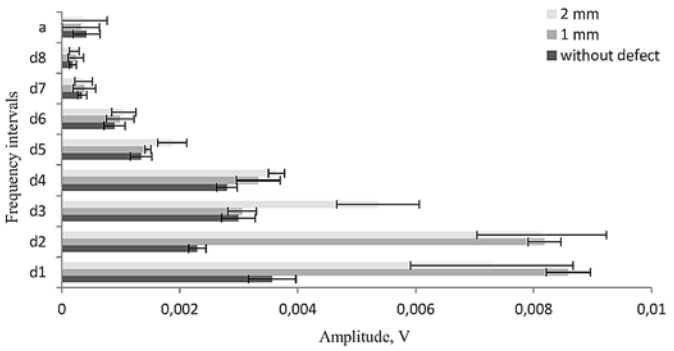


Fig. 12. The acoustic emission signal of the gear, diagnostic parameter X2 (Peak to Peak)

Table 1. Comparison of results

Results using data without wavelet decomposition, using vibro-acoustic signal (m/s^2)			
	X1	X2	X3
without defect	105.54	182.72	22.05
1 mm defect	178.83	349.99	17.48
diagnostic parameter increase %	69%	95%	21% decrease
2 mm defect	204.38	395.98	17.60
diagnostic parameter increase %	94%	117%	20% decrease
Results using proposed method in frequency range 1250–2500 Hz, using vibro-acoustic signal (m/s^2)			
	X1	X2	X3
without defect	29.76	57.93	7.52
1 mm defect	64.27	127.08	7.87
diagnostic parameter increase %	116%	119%	5%
2 mm defect	71.52	137.91	7.79
diagnostic parameter increase %	140%	138%	4%

by 1.36 times in d4 frequency range, by 1.36 times in d5 frequency range, by 1.16 times in d6 frequency range. In d1 and d2 parameter X1 decreased.

The X2 parameter (Fig. 12) has increased by 1.02 times in d3 (6250 – 12500 Hz) frequency range, after introducing a defect of 1 mm, by 1.19 times in d4 (3125 – 6250 Hz) frequency range, by 1.09 times in d5 (1562.5 – 3125 Hz) frequency range, by 1.11 times in d6 (781.3 – 1562.5 Hz) frequency range, by 3.57 times in d2 (12.5 – 25 kHz) frequency range, by 2.41 times in d1 (25 – 50 kHz) frequency range. After introducing a defect of 2 mm, X2 parameter has increased by 1.79 times in d3 frequency range, by 1.30 times in d4 frequency range, by 1.39 times in d5 frequency range, by 1.18 times in d6 frequency range. In d1 and d2 parameter X2 decreased.

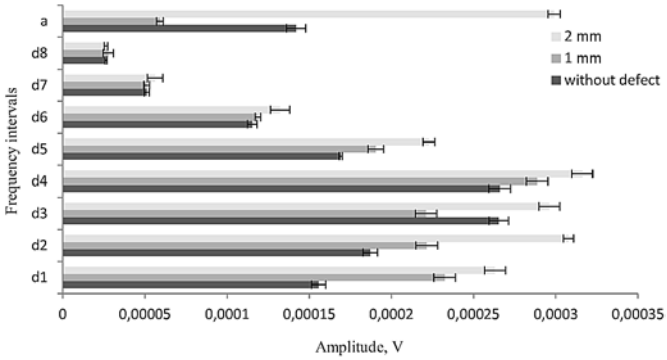


Fig. 13. The acoustic emission signal of the gear, diagnostic parameter X3 (RMS)

The X3 parameter (Fig. 13) has increased by 1.09 times in d4 frequency range, after introducing a defect of 1 mm, by 1.13 times in d5 frequency range, by 1.03 times in d6 frequency range, by 1.18 times in d2 frequency range, by 2.49 times in d1 frequency range. After introducing a defect of 2 mm, X3 parameter has increased by 1.19 times in d5 frequency range, by 1.32 times in d5 frequency range, by 1.15 times in d6 frequency range, by 1.69 times in d1 frequency range, by 1.64 times in d2 frequency range.

The analysis of the obtained results using AE indicates that defects are observed in the same frequency intervals as with the VS. However, RMS parameter, on the contrary to the VS, has increased. The parameter has also risen having processed data, and the same results have been received within mathematical modelling. While measuring the VS, 10-bit resolution was not enough, and therefore data has been distorted. Also, the application of the AE method shows that the highest signal gain can be noticed under high frequencies starting from 12500 Hz and higher. Nevertheless, confidence intervals of parameters X1 and X2 vary widely in these frequency intervals. The received data do not point to an increase in a defect; however, when the defect occurs, the signal changes considerably, and therefore this parameter can be used for detecting the defect. Same diagnostic parameters but at lower frequencies can be employed for the increase of defect.

Numerical values of diagnostic parameters determined in mathematical modelling and in the run of natural experiments vary. To reach minimum differences, further investigation on transmission measurements are required. Also, backlashes between the gears, in bearings must be established. However, this is not the focus of this paper. For diagnostic purposes, the amount of input data is minimal. Most common defects in the system are introduced, and investigation into the frequency interval where diagnostic parameters change the most is carried out. The proposed model allows assessing the obtained material, and the collected findings only prove that.

The effectiveness of the proposed diagnostic features was checked by comparing the results using data without wavelet decomposition and proposed method, in both cases diagnostic parameters where the same Peak (X1), Peak-to-Peak (X2) and RMS (X3). Using both methods, best results were obtained using Peak and Peak-to-Peak parameters, in all the cases proposed method was more sensitive to defect increase, minimal difference was 21% using X2 parameter

(gear with 2 mm defect), maximal 47% using X1 parameter (gear with 1 mm defect).

The applicability of the proposed diagnostic features was proved, and proposed method is more sensitive for gear diagnostics with selected defects.

5. Diagnostic algorithm

In this part of paper diagnostic algorithm is presented. Firstly, for condition monitoring, the technical parameters of the tested element (initial data) must be established. For testing a gear drive we need to know bearings number, information about the gear module as well as the width and number of teeth and gear ratio.

When the initial data is established, it is entered into a mathematical model and simulation takes place. After that defects are included into the system and simulation runs a few more times. Dynamical characteristics of defective and non-defective gears are compared and frequency intervals indicating the most significant variations in diagnostic parameters are defined (Fig. 14).

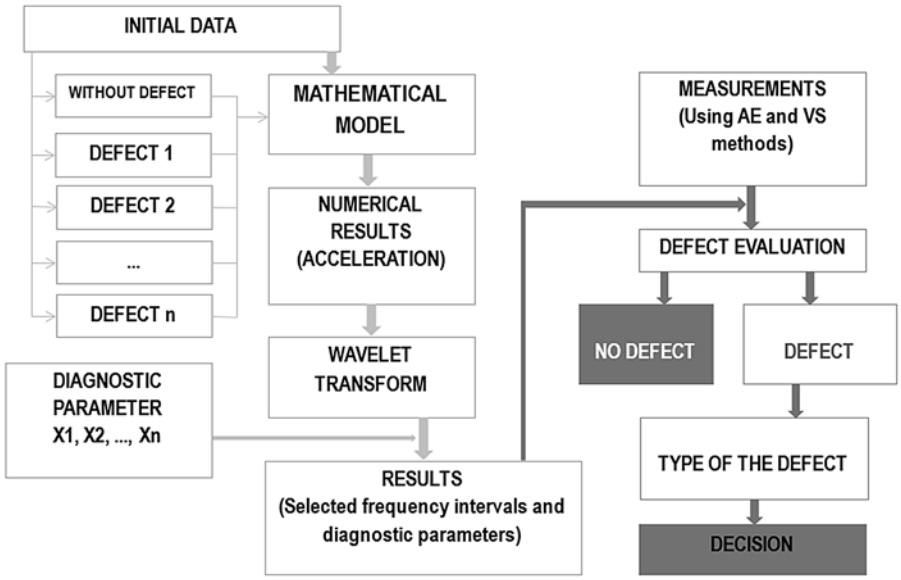


Fig. 14. Results of mathematical modelling, diagnostic parameter: a) X1 (Peak) b) X2 (Peak to Peak)

The conducted research has disclosed that diagnostic parameters Peak and Peak to Peak are the best option for spur gear condition monitoring using proposed algorithm. Upon the establishment of frequency intervals most sensitive to detecting failures, testing of the real equipment can be performed. If there are no defects in system equipment can be used, if there are defects, type of defects can be found and decision about equipment usage can be carried out.

6. Conclusions

1. Algorithm for spur gear condition monitoring has been developed. The algorithm is based on the mathematical model of the spur gear, measurements of Acoustic Emission and the Vibro-acoustic signal as well as on data processing using Wavelet transform.
2. The mathematical model of the spur gear has been presented. The model contains the evaluated backlashes between teeth, time-varying stiffness, variations in the centre distance under the working gear and shaft imbalance. The employment of the presented model assisted in analysing the dynamics of the gear with the removed top-part of the tooth. It was found that diagnostic parameters (X1 – Peak, X2 – Peak to Peak) are most sensitive to the occurrence and expansion of the defect when using the current model.
3. Frequency intervals with the most significant increase in the diagnostic parameter values have been determined. For examining the dynamic parameters of inspected defective gear, frequency interval from 1250 Hz to 5000 Hz have been applied. The measurement of the vibration signal and created mathematical model demonstrate that the most sensitive interval is the one between 1250 and 2500 Hz, and an increase in parameter X1, in this interval, was by 47 % (1 mm defect) and 46 % (2 mm defect) larger than that in the same parameter, not referring to the adopted algorithm. The AE method helps to receive good results (no errors in detecting defects are observed) within frequency intervals from 780 to 6250 Hz.
4. Proposed algorithm can be applied on gears with different ratios. There is a possibility that frequency intervals where the most significant increase in the values of diagnostic parameters can be seen, will be different from that proposed in paper, but they can be calculated easily by using mathematical model.

References

1. Amabili M, Rivola A. Dynamic analysis of spur gear pairs: steady-state response and stability of the SDOF model with time-varying meshing damping. Mech. Syst. Signal Process 1997; 11: 375–390, https://doi.org/10.1006/mssp.1996.0072.
2. Cheon G J. Nonlinear behavior. Analysis of spur gear pairs with a one-way clutch. Journal of Sound and Vibration 2007; 301: 760–776, https://doi.org/10.1016/j.jsv.2006.10.040.
3. Fakher C, Walid B, Mohamed S A, Mohamed H. Effect of spalling or tooth breakage on gear mesh stiffness and dynamic response of a one-stage spur gear transmission. European Journal of Mechanics A/Solids 2008; 27: 691–705, https://doi.org/10.1016/j.euromechsol.2007.11.005.
4. Fakhfakh T, Walha L, Louati J, Haddar M. Effect of manufacturing and assembly defects on two-stage gear system vibration. The International Journal of Advanced Manufacturing Technology 2005; 29(9): 1008–1018, https://doi.org/10.1007/s00170-004-2253-x.
5. Frolov V K, Kosarev O I. Control of gear vibrations at their source. International Applied Mechanics 2003; 39(1): 49–55, https://doi.org/10.1023/A:1023612015873.
6. He Q, Kong F, Yan R. Subspace- based gearbox condition monitoring by kernel principal component analysis. Mechanical Systems and Signal Processing 2007; 21: 1755–1772, https://doi.org/10.1016/j.ymssp.2006.07.014.
7. He S, Cho S, Singh R. Prediction of dynamic friction forces in spur gears using alternate sliding friction formulations. Journal of Sound and Vibration 2008; 309: 843–851, https://doi.org/10.1016/j.jsv.2007.06.077.

8. Jia S X, Howard I. Comparison of localized spalling and crack damage from dynamic modelling of spur gear vibrations. Mechanical Systems and Signal Processing 2006; 20: 332–349, https://doi.org/10.1016/j.ymssp.2005.02.009.
9. Kahraman A, Singh R. Non-linear dynamics of a spur gear pair. Journal of Sound and Vibration 1990; 142(1): 49-75, https://doi.org/10.1016/0022-460X(90)90582-K.
10. Kang Y, Wang C-C, Chang Y-P. Gear fault diagnosis in time domains by using Bayesian networks. Theoretical Advances and Applications of Fuzzy Logic and Soft. Series: Computing Advances in Soft Computing 2007; 42: 741–751, https://doi.org/10.1007/978-3-540-72434-6_75.
11. Kiekbusch T, Howard I. A Common Formula for the Combined Torsional Mesh Stiffness of Spur Gears. Proceedings of the 5th Australasian Congress on Applied Mechanics (ACAM 2007), Brisbane, Australia 2007; 710–716.
12. Kuang J H, Lin A D. The effect of tooth wear on the vibration spectrum of a spur gear pair. Journal of Vibration and Acoustics 2001; 123(3): 311–317, https://doi.org/10.1115/1.1379371.
13. Kuang J H, Yang Y T. An estimate of mesh stiffness and load sharing ratio of a spur gear pair. in Proceeding of ASME 6th International Power Transmission and Gearing Conference, 13–16 September 1992, Scottsdale, Arizona 1992; 1–9.
14. Lei Y, Zuo M J. Gear crack level identification based on weighted K nearest neighbour classification algorithm. Mechanical Systems and Signal Processing 2009; 23(5): 1535–1547, https://doi.org/10.1016/j.ymssp.2009.01.009.
15. Litak G, Friswell M I. Dynamics of a gear system with faults in meshing stiffness. Nonlinear Dynamics 2005; 41: 415–421, https://doi.org/10.1007/s11071-005-1398-y.
16. 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.
17. Ma H, Pang X, Feng R, Song R, Wen B. Fault features analysis of cracked gear considering the effects of the extended tooth contact. Engineering Failure Analysis 2015; 48: 105–120, https://doi.org/10.1016/j.engfailanal.2014.11.018.
18. Maliha R, Dogruer C U, Özgüven H N. Nonlinear dynamic modeling of gear-shaft-disk-bearing systems using finite elements and describing functions. Journal of Mechanical Design 2004; 126(3): 534–541, https://doi.org/10.1115/1.1711819.
19. Martin H R. Statistical moment analysis as a means of surface damage detection. Proceedings of the 7th International Modal Analysis Conference, Society for Experimental Mechanics, Schenectady, NY 1989; 1016–1021.
20. Mba D. Acoustic Emissions and monitoring bearing health. Tribology Transactions 2003; 46(3): 447–451, https://doi.org/10.1080/10402000308982649.
21. McClintic K, Lebold M, Maynard K, Byington C, Campbell R. Residual and difference feature analysis with transitional gearbox data. Proceedings of the 54th Meeting of the Society for Machinery Failure Prevention technology, Virginia Beach, VA, May 1–4 2000; 635–645.
22. Mohammeda O D, Rantatalo M, Aidanpää J O, Kumar U. Vibration signal analysis for gear fault diagnosis with various crack progression scenarios. Mechanical Systems and Signal Processing 2013; 41: 176–195, https://doi.org/10.1016/j.ymssp.2013.06.040.
23. Ozguven H N, Houser D R. Mathematical model used in gear dynamics - a review. Journal of Sound and Vibration 1988; 121: 383–411, https://doi.org/10.1016/S0022-460X(88)80365-1.
24. Qu J, Liu Z, Zuo M J, Huang H-Z. Feature selection for damage degree classification of planetary gearboxes using support vector machine. Proceedings of the Institution of Mechanical Engineers, Part C. Journal of Mechanical Engineering Science 2011; 225(9): 2250–2264, https://doi.org/10.1177/0954406211404853.
25. Sait A S, Sharaf-Eldeen Y I. A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics. Proceedings of the 29th IMAC, A Conference on Structural Dynamics 2011; 307–324, https://doi.org/10.1007/978-1-4419-9428-8_25.
26. Samanta B, Al-Balushi K R. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. Mechanical Systems and Signal Processing 2003; 17(2): 317–328, https://doi.org/10.1006/mssp.2001.1462.
27. Saxena A, Parey A, Chouksey M. Effect of shaft misalignment and friction force on time varying mesh stiffness of spur gear pair. Engineering Failure Analysis 2015; 49: 79–91, https://doi.org/10.1016/j.engfailanal.2014.12.020.
28. Skrickij V, Bogdevičius M, Junevičius R. Diagnostic features for the condition monitoring of hypoid gear utilizing the wavelet transform, Applied Acoustics 2016; 106: 51–62, https://doi.org/10.1016/j.apacoust.2015.12.018.
29. Skrickij V, Bogdevičius M. Vehicle gearbox dynamics: centre distance influence on mesh stiffness and spur gear dynamics. Transport 2010; 25(3): 278–286, https://doi.org/10.3846/transport.2010.34.
30. Staszewski W J, Worden K. Classification of faults in gearboxes – pre-processing algorithms and neural networks. Neural Computing & Applications 1997; 5(3): 160–183, https://doi.org/10.1007/BF01413861.
31. Taylor J I, Kirkland D W. The bearing analysis handbook: a practical guide for solving vibration problems in bearings. Vibration Consultant. 2004.
32. Utagawa M. Dynamic loads on spur gear teeth. The Japan Society of Mechanical Engineers 1958; 1(4): 397–403, https://doi.org/10.1299/jsme1958.1.397.
33. Vaishya M, Singh R. Sliding friction-induced non-linearity and parametric effects in gear dynamics. Journal of sound and vibration 2001; 248(4): 671–694, https://doi.org/10.1006/jsvi.2001.3818.
34. Walha L, Fakhfakh T, Haddar M. Nonlinear dynamics of a two-stage gear system with mesh stiffness fluctuation, bearing flexibility and backlash. Mechanism and Machine Theory 2009; 44(5): 1058–1069, https://doi.org/10.1016/j.mechmachtheory.2008.05.008.
35. Wang J, Li R, Peng X. Survey of nonlinear vibration of gear transmission systems. Applied Mechanics Reviews 2003; 56: 309–329, https://doi.org/10.1115/1.1555660.
36. Wojnarowski J, Onishchenko V. Tooth wear effects on spur gear dynamics. Mechanism and Machine Theory 2003; 38: 161–178, https://doi.org/10.1016/S0094-114X(02)00091-5.
37. Yang D C H, Lin J Y. Hertzian damping, tooth friction and bending elasticity in gear impact dynamics. Journal of Mechanisms, Transmissions, and Automation in Design 1987: 109(2): 189–96, https://doi.org/10.1115/1.3267437.
38. Yu J-B. Bearing performance degradation assessment using locality preserving projections. Expert Systems with Applications 2011; 38: 7440–7450, https://doi.org/10.1016/j.eswa.2010.12.079.
39. Zakrajsek J J, Townsend D P, Decker H J. An analysis of gear fault detection methods as applied to pitting fatigue failure data, The Systems engineering Approach to Mechanical Failure Prevention, Technical report, 47th Meeting of the MFPG 1993.

40. Zouari S, Maatar M, Fakhfakh T, Haddar M. Three-dimensional analyses by finite element method of a spur gear: effect of cracks in the teeth foot on the mesh stiffness. Journal of Failure Analysis and Prevention 2007; 7: 475–481, https://doi.org/10.1007/s11668-007-9078-5.

Viktor SKRICKIJ
Faculty of Transport Engineering
Vilnius Gediminas Technical University
Plytines str., 27-307 Vilnius, Lithuania

Marijonas BOGDEVIČIUS
Faculty of Transport Engineering
Vilnius Gediminas Technical University
Plytines str., 27-315 Vilnius, Lithuania

Rasa ŽYGIENĖ
Faculty of Transport Engineering
Vilnius Gediminas Technical University
Plytines str., 27-304 Vilnius, Lithuania

E-mails: viktor.skrickij@vgtu.lt, marijonas.bogdevicius@vgtu.lt, rasa.zygiene@vgtu.lt

Article citation info:
ZHANG S, SUN S, SI S, WANG P. A decision diagram based reliability evaluation method for multiple phased-mission systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 485–492, http://dx.doi.org/10.17531/ein.2017.3.20.

Shuai ZHANG
Shudong SUN
Shubin SI
Peng WANG

A DECISION DIAGRAM BASED RELIABILITY EVALUATION METHOD FOR MULTIPLE PHASED-MISSION SYSTEMS

METODA OCENY NIEZAWODNOŚCI SYSTEMÓW WIELOFAZOWYCH W OPARCIU O DIAGRAMY DECYZYJNE

The multiple phased-mission system (MPMS) exists widely in practical engineering, such as aviation, spaceflight and navigation fields. Its distinct characteristic is that the system usually performs multiple missions and each mission consists of different phases. In this paper, we mainly focus on the reliability analysis for MPMS when the components have to accomplish different missions successively. A new modeling method is proposed for MPMS analysis based on the binary decision diagram (BDD) and multi-state multi-valued decision diagram (MMDD). Through this method, different phases of missions are combined with in the whole system by certain merging rules according to the operating time of a common component. Then, the system reliability can be calculated by the common calculation methods of decision diagrams by generating the through. Finally, two case studies are implemented to demonstrate the generation of BDD/MMDD models and the evaluation of system reliability. The experiment results verified the efficiency and accuracy of the proposed modeling methods.

Keywords: *multiple phased-mission systems, binary decision diagram, multi-state multi-valued decision diagram, reliability evaluation.*

Systemy wielofazowe (Multiple Phased-Mission Systems, MPMS), t.j. systemy o wielu zadaniach okresowych są powszechnie stosowane w praktyce inżynieryjnej, np. w lotnictwie, lotach kosmicznych czy nawigacji. Cechą wyróżniającą tego typu systemy jest to, że zazwyczaj wykonują one wiele zadań, z których każde składa się z różnych faz. Głównym tematem poniższej pracy jest analiza niezawodności MPMS dla przypadków, kiedy elementy składowe muszą wykonywać różne misje jedna po drugiej. W artykule zaproponowano nową metodę modelowania dla celów analizy MPMS opartą na koncepcji binarnego diagramu decyzyjnego (binary decision diagram, BDD) oraz wielostanowego wielowartościowego diagramu decyzyjnego (multi-state multi-valued decision diagram, MMDD). Metoda ta polega na łączeniu różnych faz misji w obrębie systemu za pomocą pewnych reguł łączenia wedle czasu pracy wspólnego elementu składowego. Pozwala to na obliczanie niezawodności systemu za pomocą powszechnie stosowanych metod diagramów decyzyjnych poprzez generowanie drzew błędów. W pracy zaprezentowano dwa studia przypadku, które pokazują, w jaki sposób generuje się modele BDD/MMDD oraz ocenia niezawodność systemu. Wyniki eksperymentów wykazały wydajność oraz trafność proponowanych metod modelowania.

Słowa kluczowe: *systemy wielofazowe, binarny diagram decyzyjny, wielostanowy wielowartościowy diagram decyzyjny, ocena niezawodności.*

Notations

a, b, c, A_r	the component in MPMS
r	the ID of a component
w	the total number of components in the system
x_{A_r}	state variable of component A_r
m, n	state of component
i, j	phase of the system
$x_{A_{ri}}$	state variable of component A_r in phase i
P, Q	mission of the system
P_i, Q_i	phases of the mission P and mission Q
t_{A_r}	the component A_r which works on the two missions' time nodes
G,H	Boolean functions
$\text{index}(x_{A_r})$	position of x_{A_r} in the propagation order of all BDD variables
F_i	logical expression of phase i

Acronyms

PMS	phased-mission system
MPMS	multiple phased-mission system
BDD	binary decision diagram
MFTA	multi-state fault tree analysis
MMDD	multi-state multi-valued decision diagram
MSS	multi-state system
DAG	directed acyclic graph
PDO	phase-dependent operation
ite	if-then-else

1. Introduction

Phased-mission systems (PMS) are very common in practical engineering, where the mission of system usually consists of multiple, consecutive, and non-overlapping phases in operation [12, 20, 21]. A simple example is that the phases of car-driving mission include

start, acceleration, deceleration, and stop. During each phase, the system has to complete the specific task and may be subject to different stresses and environmental conditions as well as different reliability requirements [12]. Moreover, the system's functioning principle of different phases may change, and hence it is necessary to establish distinct models for each phase.

Accurate reliability analysis of PMS must consider the statistical dependencies of components across different phases, as well as the dynamics of system configurations, success criteria, and component behavior. In the previous study, researchers mainly focused on binary reliability models for PMS. Park and Yoo [11] introduced an iterative Lagrange technique to maximize the mission reliability of PMS by apportioning subsystem reliabilities according to multiple resource constraints. Dugan [2] proposed an automated analysis method of PMS based on the discrete-state continuous-time Markov model. Kim and Park [4] put forward three cases, whose phase durations are deterministic, random variables exponential distribution, to compute the mission reliability based on Markov model. Somani and Trivedi [13] proposed a Boolean algebraic method to analyze PMS reliability, and the failure criterion in each phase can be expressed as a fault tree. Ma and Trivedi [8] described an efficient Boolean algebraic algorithm which combines the fault trees of all the phases into a single fault tree with repeated events. Zang et al [24] established a method based on binary decision diagram (BDD) to analyze the reliability of PMS. Jung et al [3] proposed a BDD algorithm for coherent fault tree, where the truncated if-then-else (ite) connectives and subsuming could be performed in the progress of the BDD structure construction.

Recently, more and more researchers have been concentrated on multi-state systems (MSS) and multi-state PMS. Tang and Dugan [18] built the dependence-BDD for reliability analysis of PMS with multi-mode failures by applying dependence algebra. Xing and Dai [22] proposed a new modeling approach called multi-state multi-valued decision diagrams (MMDD) for the analysis of multi-state systems. Shrestha and Xing [14-16] introduced reliability analysis of multi-state PMS with unordered and ordered states, and used MMDD to analyze the importance of components. Levitin and Xing [5,6] introduced a recursive algorithm based on conditional probability and an efficient recursive formula based on the branch and bound method for reliability evaluation of non-repairable PMS. Xing and Amari [23] put forward an efficient method to evaluate the reliability of k-out-of-n systems with identical components subject to phased-mission requirements and imperfect fault coverage. Wang and Xing [19] established an algorithm for competing failure analysis in PMS with functional dependence in one of the phases. Zang and Bai [25] proposed a mathematical model for success probability analysis of PMS based on minimal path set and system state analysis methods. Mo and Xing [9,10] built a new analytical method based on multi-valued decision diagrams for reliability analysis of non-repairable PMS with multi-mode failures. Li and Tao [7] combined the Bayesian networks with event tree and fault tree analysis to analyze PMS based on conditional probability by giving expression of the phase-dependency.

Multiple phased-mission systems (MPMS) have been applied in a wide range of engineering fields, where a system consists of multiple missions. The state of the component at the end of a mission will be the beginning state of the same component in the next mission. In MPMS, each mission also consists of multiple, consecutive, and non-overlapping phases which are accomplished in sequence. For example, the operational process of landing gear involves two missions: take-off and landing. The take-off mission involves speed skating, lifting, and climbing phases. And the landing mission involves landing gear drop-down, level flight, drift down, and skating phases. The landing gear system needs to complete both two missions for success flight. Compared with PMS, the analysis of MPMS is more difficult because a component may work during two missions in sequence. In PMS, for the component working in different phases, all the phases can be merged as one by existing algorithms. But in the MPMS, it is usually assumed that a component have to work in two missions in sequence. System structure and the environmental conditions will make the state of components more complex. So we need generate some new phases for the common component which works on the two missions' time nodes, and then combine all the phases.

The remainder of this paper is organized as follows. Section 2 presents the basic concept and phase-dependent operation algorithm of BDD and MMDD respectively. Section 3 describes the reliability evaluation methods of MPMS based on BDD and MMDD. Two examples are illustrated in Section 4 to show the efficiency and accuracy of the proposed modeling methods. Section 5 gives conclusion and points out the future work.

2. Methodologies

2.1. Basic concept of BDD

BDD is a rooted, directed acyclic graph representation of a Boolean expression based on Shannon decomposition rule [1]. It has two sink nodes (outputs), labeled as '1' and '0', which represent a binary-state system being either operational or failed. Let $A_r (r = 1,2,...,w)$ be the component. Let w denote the total number of components in the system. The two states of component A_r represented by a Boolean variable, denoted by x_{A_r}. Each Boolean variable x_{A_r} can be represented using the if-then-else (ite) format as $ite(x_{A_r},1,0)$. In general, the (ite) format for expressing Boolean expressions F (representing the system state structure function) in variable x_{A_r} based on Shannon's decomposition is: $F = x_{A_r} \cdot F_{x_{A_r}=1} + \overline{x_{A_r}} \cdot F_{x_{A_r}=0} = ite(x_{A_r}, F_{x_{A_r}=1}, F_{x_{A_r}=0})$. In practical engineering, non-sink node usually corresponds to the component's state. By traversing the BDD's all paths with each path pointing to sink node '1', the probability of occurrence of the system can be calculated.

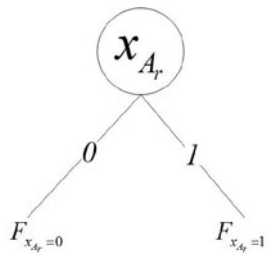


Fig. 1. Binary decision diagram

Each non-sink node in BDD usually has two outgoing edges, called 0-edge and 1-edge, respectively. Supposing there are two sub-BDD models G and H, then they could be encoded with the Boolean expression in the ite format, as:

$$G = x_{A_r} \cdot G_{x_{A_r}=1} + \overline{x_{A_r}} \cdot G_{x_{A_r}=0} = ite(x_{A_r}, G_{x_{A_r}=1}, G_{x_{A_r}=0}) = ite(x_{A_r}, G_1, G_0),$$

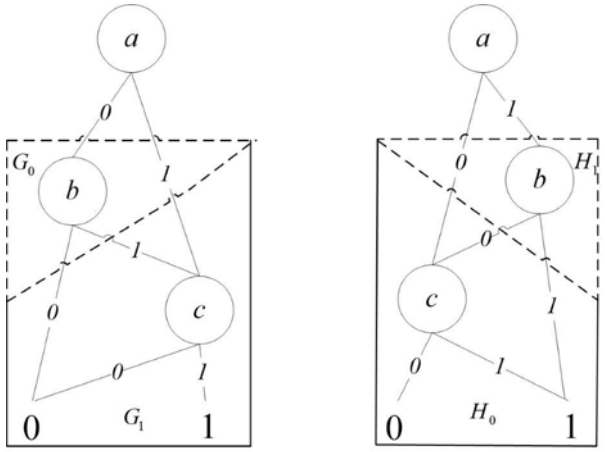
$$H = y_{A_r} \cdot H_{y_{A_r}=1} + \overline{y_{A_r}} \cdot H_{y_{A_r}=0} = ite(y_{A_r}, H_{y_{A_r}=1}, H_{y_{A_r}=0}) = ite(y_{A_r}, H_1, H_0).$$

Phased-mission systems (PMS) are systems in which multiple non-overlapping phases of tasks are accomplished in sequence for a successful mission. To combine different phases, the operation rules for combing two sub-BDD models G and H are as:

$$G \Diamond H = \text{ite}(x_{A_r}, G_1, G_0) \Diamond \text{ite}(y_{A_r}, H_1, H_0)$$
$$= \begin{cases} \text{ite}(x_{A_r}, G_1 \Diamond H_1, G_0 \Diamond H_0) & \text{index}(x_{A_r}) = \text{index}(y_{A_r}) \\ \text{ite}(x_{A_r}, G_1 \Diamond H, G_0 \Diamond H) & \text{index}(x_{A_r}) < \text{index}(y_{A_r}) \\ \text{ite}(y_{A_r}, G \Diamond H_1, G \Diamond H_0) & \text{index}(x_{A_r}) > \text{index}(y_{A_r}) \end{cases}, \quad (1)$$

where the symbol $\Diamond$ represents a logic operation (AND or OR) between two sub-BDD models, the index( ) is assigned to each variable to indicate its position in the propagation order of all BDD variables. For example, $\text{index}(x_{A_r}) < \text{index}(y_{A_r})$ implies that the position of the y_{A_r} is behind the position of the x_{A_r} in the order.

To clearly explain the operation rules in equation (1), the detailed examples of two sub-BDD models G and H are shown in Fig.2[24]. For sub-BDD models G in Fig.2 (a), we know that $G = a \cdot G_1 + \bar{a} \cdot G_0$. Since $G_1 = c \cdot 1 + \bar{c} \cdot 0 = c$ and $G_0 = b \cdot G_1 + \bar{b} \cdot 0 = b \cdot c$, then we can get $G = a \cdot c + \bar{a} \cdot b \cdot c = a \cdot c + b \cdot c$. For sub-BDD models H in Fig.2 (b), we know that $H = a \cdot H_1 + \bar{a} \cdot H_0$. Since $H_0 = c \cdot 1 + \bar{c} \cdot 0 = c$ and $H_1 = b \cdot 1 + \bar{b} \cdot H_0 = b + \bar{b} \cdot c = b + c$, then we can get $H = a \cdot (b + c) + \bar{a} \cdot c = a \cdot b + a \cdot c + \bar{a} \cdot c = a \cdot b + c$.



(a) $G = a \cdot c + b \cdot c$ (b) $H = a \cdot b + c$

Fig. 2. The detailed examples of two sub-BDD models

If G and H are connected with "OR" operator, then the combination process of two sub-BDD models in Fig.2 is shown in Fig.3.

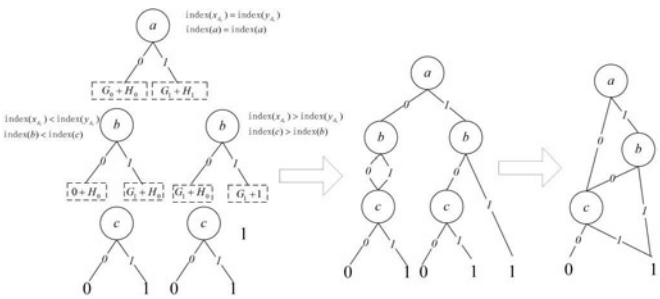


Fig. 3. Combination of two sub-BDD G and H with "OR" operation

Generally, the combination process of sub-BDD models could be concluded as follows:

(1) Compare the two sub-BDD models, it is clear that $\text{index}(a) = \text{index}(a)$. According to the rules of equation (1), we have $\text{ite}(a, G_1 \Diamond H_1, G_0 \Diamond H_0)$.

(2) Compare the G_0 and H_0, we know that $\text{index}(b) < \text{index}(c)$. According to the rules of equation (1), we can get $\text{ite}(b, G_1 \Diamond H_0, 0 \Diamond H_0)$.

(3) Compare the G_1 and H_1, we have $\text{index}(c) > \text{index}(b)$. Accroding to the rules of equation (1), we can get $\text{ite}(b, G_1 \Diamond 1, G_1 \Diamond H_0)$

Simplify the process is as follows:

(1) Because the results of $0 + H_0$ and $G_1 + H_0$ are the same, the node b 0-edge and 1-edge all point to the node c. So the node b at left can be removed.

(2) Because the 0-edge of node a is point to the same sub-tree as the 0-edge of node b at right. One of the two same sub-tree can be reduced.

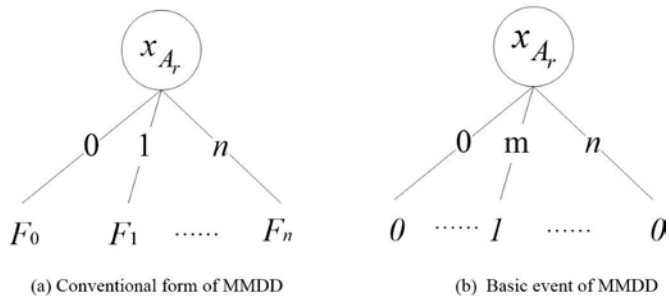
2.2. Basic concept of MMDD

The MMDD is a multi-state extended form of BDD [16]. It is a multi-valued logic structure for the natural representation of the MSS and is widely used in MSS reliability analysis [17]. The nodes MMDD are also divided into two types: sink nodes and non-sink nodes. MMDD only has two sink nodes, labeled '1' and '0', which indicate that the system is either in state '1' or in state '0'. Non-sink node in MMDD can have more than two edges where each edge represents a possible state of the components.

According to [17], Logical expression F in MMDD can be represented as follows:

$$\begin{aligned} F &= A_0 \cdot F_{x_{A_0}=0} + A_1 \cdot F_{x_{A_1}=1} + \ldots + A_n \cdot F_{x_{A_n}=n} \\ &= \text{case}(A_r, F_{x_{A_0}=0}, F_{x_{A_1}=1}, \ldots F_{x_{A_n}=n}) \\ &= \text{case}(A_r, F_0, F_1, \ldots F_n) \end{aligned} \quad , \tag{2}$$

Each non-sink node is associated with a multi valued state variable x_{A_r}, and $x_{A_r} = m$ means that the component A_r is in state m. The F_m can take one of two values: "1" or "0", indicating that F is in or not in state m(m=0,1,2,…,n) respectively. The non-sink note x_{A_r} has $(n+1)$ possible states and can be in a particular state at a specific time. So the logical expression of F has $(n+1)$ possible values. For example, when $x_{A_r} = m$, $F_0 = 0, F_1 = 0, \ldots, F_m = 1, \ldots, F_n = 0$. When sink node x_{A_r} is in state m, the value of F is '1'; otherwise the value is '0', as shown in Fig. 4.



(a) Conventional form of MMDD (b) Basic event of MMDD

Fig. 4. MMDD of node x_{A_r}.

2.3. Phase-dependent operation of decision diagram

In 1999, Zang et al [24] published a paper about the application of BDD for phased-mission systems and derived a special phased-dependent operation (PDO) as in equations (3) and (4). Let component A_r be used in both phase i and j, $i < j$. Using ite format, F_i, F_j express the Boolean expressions of F is in phase i and j, while $x_{A_{ri}}$ denoted the state variable of component A_r in phase i. Then we have: $G_1 = (F_i)_{x_{A_{ri}}=1}, G_0 = (F_i)_{x_{A_{ri}}=0}, H_1 = (F_j)_{x_{A_{rj}}=1}, H_0 = (F_j)_{x_{A_{rj}}=0}$.

$$\begin{aligned} F_i &= ite\left[x_{A_{ri}}, (F_i)_{x_{A_{ri}}=1}, (F_i)_{x_{A_{ri}}=0}\right] = ite\left(x_{A_{ri}}, G_1, G_0\right) \\ F_j &= ite\left[x_{A_{rj}}, (F_j)_{x_{A_{rj}}=1}, (F_j)_{x_{A_{rj}}=0}\right] = ite\left(x_{A_{rj}}, H_1, H_0\right) \end{aligned} \tag{3}$$

$$F_i \Diamond F_j = \begin{cases} ite\left(x_{A_{ri}}, G_1 \Diamond F, G_0 \Diamond F_0\right) ForwardPDO \\ ite\left(x_{A_{rj}}, G_1 \Diamond F_1, G \Diamond F_0\right) BackwardPDO \end{cases} \tag{4}$$

Because BDD modeling process depends on the order of the variables, there are two types of ordering methods: forward PDO and backward PDO. For the forward PDO, the variable order is the same as the phase order $(x_{A_{r1}}, x_{A_{r2}} ..., x_{A_{rs}})$. In the backward PDO, the variable order is the reverse of the phase order $(x_{A_{rs}}, x_{A_{r(s-1)}} ..., x_{A_{r1}})$. For combined operations, the same component belongs to two sub-BDDs but in different phases.

To deal with the MPMS problems, we derive a new MMDD operation for Phase Algebra in this paper based on the results of [24]. Similarly, two types of ordering methods are considered: forward PDO and backward PDO. Let component A appear in both phase i and j, $i < j$, then we have:

$$F_i = case\left[x_{A_{ri}}, (F_i)_{x_{A_{ri}}=0}, (F_i)_{x_{A_{ri}}=1} ... (F_i)_{x_{A_{ri}}=n}\right] = case\left(x_{A_{ri}}, G_0, G_1, ... G_n\right) = G$$
$$F_j = case\left[x_{A_{rj}}, (F_j)_{x_{A_{rj}}=0}, (F_j)_{x_{A_{rj}}=1} ... (F_j)_{x_{A_{rj}}=n}\right] = case\left(x_{A_{rj}}, H_0, H_1, ... H_n\right) = H, \tag{5}$$

$$F_i \Diamond F_j = \begin{cases} case\left(x_{A_{ri}}, G_0 \Diamond H_0, G_1 \Diamond H, G_2 \Diamond H, ... G_n \Diamond H\right) ForwardPDO \\ case\left(x_{A_{rj}}, G \Diamond H_0, G \Diamond H_1, G \Diamond H_2, ... G_n \Diamond H_n\right) BackwardPDO \end{cases}. \tag{6}$$

(i) For the forward PDO,

If x_{A_r} is failed in phase i and further it is irreparable, then it keeps failed in phase j, i.e, $x_{A_{ri}} = 0$ implies $x_{A_{rj}} = 0$.

$$\begin{aligned} F_i \Diamond F_j &= case\left(x_{A_{ri}}, G_0, G_1, ... G_n\right) \Diamond case\left(x_{A_{rj}}, H_0, H_1, ... H_n\right) \\ &= case\left[x_{A_{ri}}, (F_i \Diamond F_j)_{x_{A_{ri}}=0}, (F_i \Diamond F_j)_{x_{A_{ri}}=1} ... (F_i \Diamond F_j)_{x_{A_{ri}}=n}\right] \\ &= case\left[x_{A_{ri}}, (F_i)_{x_{A_{ri}}=0} \Diamond (F_j)_{x_{A_{ri}}=0}, (F_i)_{x_{A_{ri}}=m} \Diamond (F_j)_{x_{A_{ri}}=m}, (F_i)_{x_{A_{ri}}=n} \Diamond (F_j)_{x_{A_{ri}}=n}\right] \\ &= case\left(x_{A_{ri}}, G_0 \Diamond H_0, G_1 \Diamond H, G_2 \Diamond H, ... G_n \Diamond H\right) \end{aligned} \tag{7}$$

This derivation uses the equation:

$$(F_j)_{x_{A_{ri}}=m} = F_j = H \tag{8}$$

Since $x_{A_{ri}} = m$ is not relevant to H.

(ii) For the backward PDO,

If x_{A_r} is operational in phase j, then it must be operational in phase i, i.e, $x_{A_{rj}} = n$ implies $x_{A_{ri}} = n$.

$$\begin{aligned} F_i \Diamond F_j &= case\left(x_{A_{ri}}, G_0, G_1, ... G_n\right) \Diamond case\left(x_{A_{rj}}, H_0, H_1, ... H_n\right) \\ &= case\left[x_{A_{rj}}, (F_i \Diamond F_j)_{x_{A_{rj}}=0}, (F_i \Diamond F_j)_{x_{A_{rj}}=1} ... (F_i \Diamond F_j)_{x_{A_{rj}}=n}\right] \\ &= case\left[x_{A_{rj}}, (F_i)_{x_{A_{rj}}=0} \Diamond (F_j)_{x_{A_{rj}}=0}, (F_i)_{x_{A_{rj}}=m} \Diamond (F_j)_{x_{A_{rj}}=m}, (F_i)_{x_{A_{rj}}=n} \Diamond (F_j)_{x} \right. \\ &= case\left(x_{A_{rj}}, G \Diamond H_0, G \Diamond H_1, G \Diamond H_2, ... G_n \Diamond H_n\right) \end{aligned} \tag{9}$$

This derivation uses the equation:

$$(F_i)_{x_{A_{rj}}=m} = F_i = G \tag{10}$$

Since $x_{A_{rj}} = m$ is not relevant to G.

In the forward PDO, index($x_{A_{ri}}$)<index($x_{A_{rj}}$) when phase $i < j$. The new MMDD node of the combined sub-MMDD is $x_{A_{ri}}$. The 0-edge of node $x_{A_{ri}}$ is generated, and the operation is applied to G_0 and H_0. In order to generate the m-edge of node $x_{A_{ri}}$ in a combined MMDD, the operation is applied to $G_m\left(m = 1, 2, ... n\right)$ and the other sub-MMDD model H together. In the backward PDO, index($x_{A_{rj}}$)<index($x_{A_{ri}}$) when phase $i < j$. The new MMDD node of the combined sub-MMDDs is $x_{A_{rj}}$. The n-edge of node $x_{A_{rj}}$ is generated, and the operation is applied to G_n and H_n. In order to generate the m-edge of node $x_{A_{rj}}$ in a combined MMDD, the operation is applied to $H_m\left(m = 0, 1, ... n-1\right)$ and the other sub-MMDD model G together.

3. Reliability evaluation methods based on decision diagram

The proposed BDD and MMDD methods for MPMS are established according to the following assumptions: (1) Each mission consists of multiple non-overlapping phases; (2) The component completes its missions in sequence. In binary-state MPMS, both a system and its components have two and only two states: functioning or failed, which are labeled as '1' and '0', respectively. In multi-state MPMS, a system has two and only two states, while the components may have more than two states.

In order to evaluate the reliability of MPMS by the BDD and MMDD methods, we need to build the system structure function for each phase. The logical expression (representing the failure of the system in the phase) can be obtained by complementing the system structure function. The graphical representation of the logic expressions in terms of logic AND/OR gates gives the fault tree model for each phase. Based on the generated fault tree models, the proposed BDD-based analysis and the MMDD-based analysis can be performed in the following four steps:

Step 1: New Phase Generation. One component participates in multiple missions consecutively. We can see that a new system consists of two missions and there's a common component which works on two missions sequentially. The new system is divided into many phases by the common component which works on the two missions' time nodes.

Step 2: Single-Mission BDD/MMDD Generation. Traditional method can be used for the generation of BDD model for each phase. In particular, equation (1) is applied to generate BDD based on the fault tree. The MMDD model is generated according to equation (2).

Step 3: Multiple Missions of BDD/MMDD Merged for the Same Phase. Based on the result from the Step 1, we need to merge the two missions of BDD/MMDD in the same phase. Each phase of the BDD/MMDD is generated by performing the logic OR operation of the single phase BDDs/MMDDs generated in Step 2. In a binary-system, equation (4) is applied when operation is performed on two variables of different elements. In a multi-state system, equation (6) is applied when operation is performed on two variables of different elements.

Step 4: Generation of BDD/MMDD for MPMS. In this step, the entire MPMS is generated by performing the logic OR operation on all the merged BDD/MMDDs generated in Step 3. In a binary-system, equation (4) is applied when operation is performed on two nodes which belong to the same component but in different phases. In a multi-state system, equation (6) is applied when operation is performed on two nodes that belong to the same component but in different phases.

4. Illustrative examples

4.1 BDD example

An example is presented to illustrate the application of modeling method based on BDD for a system with one component being engaged in two missions: mission P and mission Q. Mission P needs two components, A_1 and A_2. Mission Q needs three components, A_1, A_3, and A_4. During different time periods, component A_1 participates in mission P and mission Q.

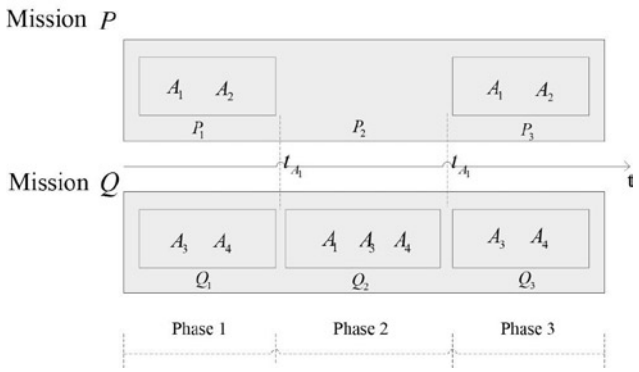


Fig. 5. System structure of BDD example

Step 1: In the system consisting of mission P and mission Q, the common component A_1 works in both missions sequentially. The new system is divided into three phases by the common component A_1 which works on the two missions' time nodes. The P_i and Q_i denoted the phases of the mission P and mission Q. The first phase consists of P_1 and Q_1. The second phase consists of Q_2. The third phase consists of P_3 and Q_3.

The input parameters of each component are shown in Table 1, which shows the computed conditional reliability for each component at each phase, and all the components fail exponentially with constant failure rate.

Table 1. Input parameters about each component.

component	Phase 1	Phase 2	Phase 3
A_1	0.968507	0.980199	0.973215
A_2	0.984127	0.980172	0.965432
A_3	0.974321	0.983543	0.981240
A_4	0.995432	0.984201	0.971205

Step 2: A single-mission single-phase BDD is generated. In a binary-state system, both the system and its components have only two states: '1' and '0', which represents the binary-state system and its components' state: either operational or failed.

In mission P of the phase 1, there are two components: A_1 and A_2. Each component has two states (0, 1). When A_1 and A_2 are in state '1', the system is normal. Fig. 6 shows the BDD model about P_1, where A_{ri} denoted the component A_r in phase i.

In mission Q of the phase 1, there are two components: A_3 and A_4. Each component has two states (0, 1). When both A_3 and A_4 are in state '1', the system is normal. Fig. 7 shows the BDD model about Q_1.

In mission Q of the phase 2, three are three components: A_1, A_3, A_4 and each component has two states (0, 1). When A_1 is in state '0', A_4 in state '1', the system is normal. When A_1 is in state '1', A_3 in state '0', A_4 in state '1', the system is normal. When A_1 is in state '1', A_3 in the state '1', the system is also normal. Fig. 8 shows the BDD model about Q_2.

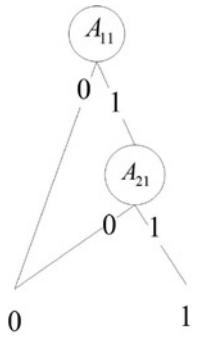


Fig. 6. The BDD model of the phase 1 of mission P

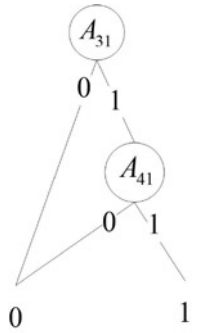


Fig. 7. The BDD model of the phase 1 of the mission Q

In mission P of the phase 3, there are two components: A_1 and A_2, and each component has two states (0, 1). When A_1 is in state '1'; or A_1 in state '0', A_2 in state '1', the system is normal. Fig. 9 shows the BDD model about P_3.

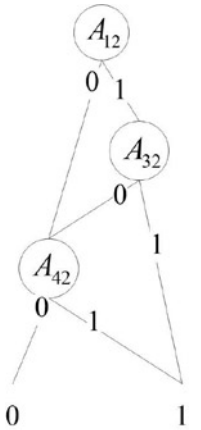


Fig. 8. The BDD model of the phase 2 of the mission Q.

In mission Q of the phase 3, there are two components: A_3 and A_4, and each component has two states (0, 1). When both component A_3 and A_4 are in state '1', the system is normal. Fig. 10 shows BDD model about Q_3.

Step 3: BDDs of two sub-missions merged for the same phase. By applying equation (4) with the order of $index(A_{11}) < index(A_{21}) < index(A_{31}) < index(A_{41}) < index(A_{12}) < index(A_{32}) < index(A_{42}) < index(A_{13}) < index(A_{23}) < index(A_{33}) < index(A_{43})$, the new BDD for each phase is presented in Fig.11.

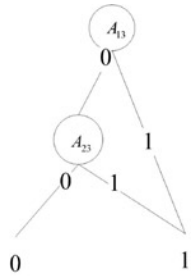


Fig. 9. The BDD model of the phase 3 of the mission P

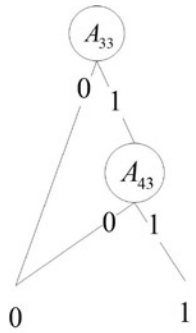


Fig. 10. The BDD model of the phase 3 of the mission Q

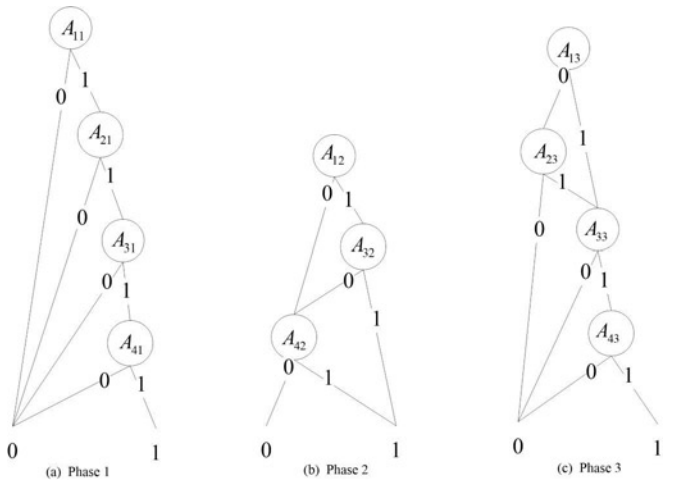


Fig. 11. Merged BDD models of each phase

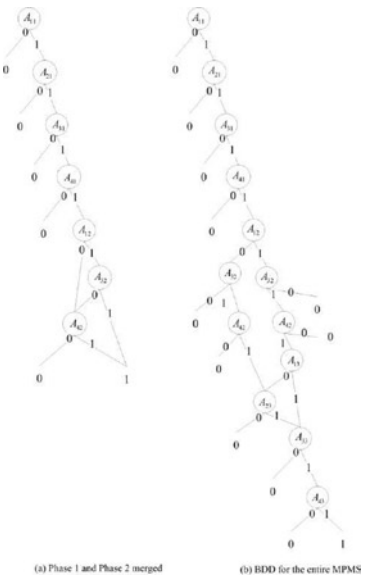


Fig. 12. Merged BDD model

Step 4: Generation of BDD for MPMS. Perform logic OR operation to combine the merged BDDs of three phases by applying equation (4) as shown in Fig.12.

Finally, according to the built BDD for entire MPMS, the overall system reliability is 0.829358.

4.2. MMDD example

We assume that a multi-state system is composed of two missions: mission P and mission Q. The step-by-step analysis of the multi-state MPMS is given as follows.

Step 1: New phase generation. Depending on the mission time

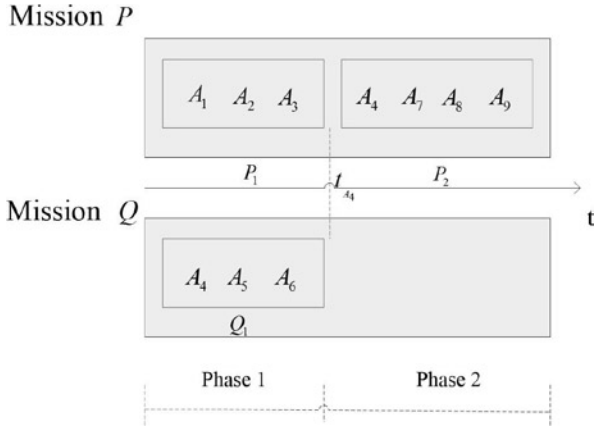


Fig. 13. System structure of MMDD example

Table 2. Input parameters of components.

component	Phase 1	Phase 2
A_1	0.0101(1), 0.982541(2)	-
A_2	0.984127	-
A_3	0.980789	-
A_4	0.986243	0.978568
A_5	0.980741	-
A_6	0.0101(1), 0.9804(2)	-
A_7	-	0.0131(1), 0.9769(2)
A_8	-	0.987562
A_9	-	0.986524

node of component A_4, mission Q has one phase: Q_1. Mission P can be divided into two phases: P_1 and P_2. In mission Q, phase Q_1 consists of three components (A_4, A_5, A_6) and is finished at time $t_{A_4}(Q)$. In mission P, phase P_1 consists of three components (A_1, A_2, A_3). At time $t_{A_4}(Q)$, component A_4 completes its task in mission Q and starts its task in mission P, meaning that phase P_2 consists of four components (A_7, A_8, A_9, A_4).The new system will be divided into two phases: 1) Phase 1 includes P_1 and Q_1; 2) phase 2 includes P_2. For the entire system, component A_4 involves in two missions during different time periods. Table 2 shows the computed conditional reliability for each element at each phase.

Step 2: Built MMDD for each phase. In phase 1 of mission P, the Component A_1 has three states (0, 1, and 2); component A_2 and A_3 have two states (0, 1). When component A_1 is in state '1', A_3 is in state '1', the system is normal. When component A_1 is in state '2' and A_2 in state '1', the system is normal. When component A_1 is in

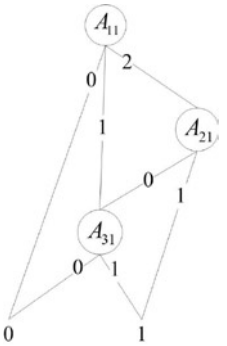


Fig. 14. The MMDD model of the phase 1 of the mission P

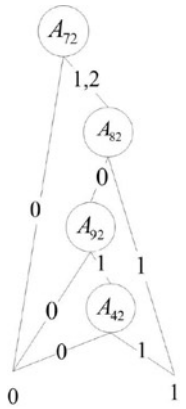


Fig. 15. The MMDD model of the phase 2 of the mission P

state '2' and A_3 in state '1', the system is also normal. Fig.14 shows the MMDD model for P_1.

In phase 2 of mission P, there are four components, when component A_4 is in state '1', A_9 is in state '1', and A_7 is in state '1' or '2', the system is normal. If component A_8 is in state '1', and A_7 is in state '1' or '2', the system is also normal. Fig.15 shows the MMDD model for P_2.

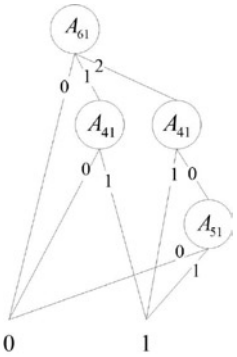


Fig. 16. The MMDD model of the phase 1 of the mission Q

In phase 1 of mission Q, there are three components: A_4, A_5, and A_6. Component A_4 and A_5 have two states (0, 1), and component A_6 has three states (0, 1, and 2). When component A_6 is in state '1' or '2', A_4 is in state '1', the system is normal. When component A_6 is in state '2' and A_5 in state '1', the system is also normal. Fig.16 shows the MMDD model for Q_1.

Step 3: Two missions MMDD merged for the same phase. Applying equation (6), and using the order

$$index(A_{11}) < index(A_{21}) < index(A_{31}) < index(A_{61}) < index(A_{41}) < index(A_{51}) < index(A_{72}) < index(A_{82}) < index(A_{92}) < index(A_{42}),$$

the new MMDD for each phase is presented in Figs.17- 18.

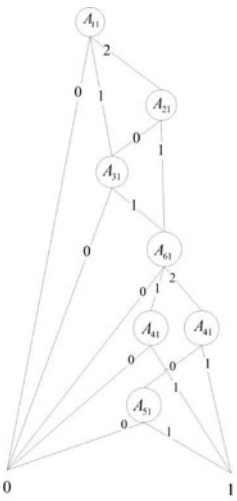


Fig. 17. Merged MMDD models of the phase 1

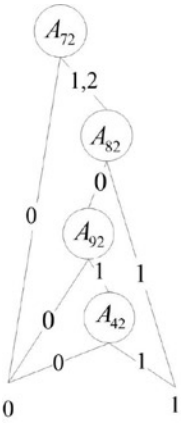


Fig. 18. Merged MMDD models of the phase 2

Step 4: Generate MMDD of entire MPMS. By performing logic OR operation to combine the MMDD of first and second phases, we obtain the final MMDD of MPMS. This is achieved by applying equation (6) on the example system, as shown in Fig.19.

Finally, according to the merged MMDD for the entire MPMS, the overall system reliability is 0.971929.

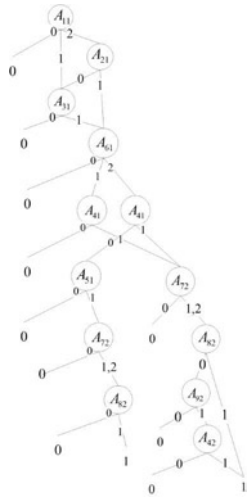


Fig. 19. MMDD for the entire MPMS

5. Conclusions

In this paper, we proposed an analytical method based on BDD and MMDD to analyze the reliability for MPMS, where the components are engaged in multiple phased-missions sequentially. A four-step procedure is proposed to generate the BDD/MMDD model for obtaining the reliability value of the MPMS. In the MMDD modeling process, the merger regulation was proved. And two examples are implemented to prove the feasibility of the proposed methods.

Acknowledgment

This work is supported by the Fundamental Research Funds for the Central Universities of China (No. 3102014JCS05013), the Aeronautical Science Foundation of China (No. 2014ZG53080), and the National Natural Science Foundation of China (No.71271170).

References

1. Bryant R. E. Graph-based algorithms for Boolean function manipulation. IEEE Transactions on Computers 1986; 35(8): 677-91, https://doi.org/10.1109/TC.1986.1676819.
2. Dugan J. B. Automated Analysis of Phased-Mission Reliability. IEEE Transactions on Reliability 1991; 40(1): 45-55, https://doi.org/10.1109/24.75332.
3. Jung W. S., Han S. H., Ha J. A fast BDD algorithm for large coherent fault trees analysis. Reliability Engineering & System Safety 2004; 83: 369-374, https://doi.org/10.1016/j.ress.2003.10.009.
4. Kim K., Park K.S. Phased-mission system reliability under Markov environment. IEEE Transactions on Reliability 1994; 43(2): 301-309Jun.
5. Levitin G., Xing L. D. Recursive algorithm for reliability evaluation of non-repairable phased mission systems with binary elements. IEEE Transactions on Reliability 2012; 61(2): 533-542, https://doi.org/10.1109/TR.2012.2192060.
6. Levitin G., Xing L. D. Reliability of non-repairable phased-mission systems with propagated failures. Reliability Engineering & System Safety 2013; 119: 218-228, https://doi.org/10.1016/j.ress.2013.06.005.
7. Li X .T.,Tao L. M. A Bayesian networks approach for event tree time-dependency analysis on phased-mission system. Eksploatacja i Niezawodnosc - Maintenance and Relaibility 2015; 17(2): 273-281, https://doi.org/10.17531/ein.2015.2.15.
8. Ma Y.,Trivedi K. S. An algorithm for reliability analysis of phased-mission systems. Reliability Engineering & System Safety 1999; 66: 157-170, https://doi.org/10.1016/S0951-8320(99)00033-2.
9. Mo Y. C., Xing L. D. MDD-Based method for efficient analysis on phased-mission systems with multimode failures. IEEE Transactions on systems man cybernetics-systems 2014; 44(6): 757-769, https://doi.org/10.1109/TSMC.2013.2277692.
10. Mo Y. C., Xing L. D. A multiple-valued decision diagram based method for efficient reliability analysis of non-repairable phased-mission systems. IEEE Transactions on Reliability2014; 63(1): 320-330, https://doi.org/10.1109/TR.2014.2299497.
11. Park K. S., Yoo Y. K. Reliability apportionment for phased-mission oriented systems. Reliability Engineering & System Safety 1990; 27: 357-364, https://doi.org/10.1016/0951-8320(90)90006-9.
12. Somani A. K., Ritcey J. A., Au S. H. L. Computationally efficient phased-mission reliability analysis for systems with variable configurations. IEEE Transactions on Reliability 1992; 41(4): 504-511, https://doi.org/10.1109/24.249576.
13. Somani A .K., Trivedi K. S. Phased-mission system analysis using Boolean algebraic methods. Proc. Association for Computing Machinery Sigmetrics Conference Meas. Model. Comput. Syst. 1994: 98-107.
14. Shrestha A,, Xing L. D. Reliability Analysis of Multistate Phased-Mission Systems With Unordered and Ordered States. IEEE Transactions on Systems and Cybernetics 2011; 41(4): 625-636, https://doi.org/10.1109/TSMCA.2010.2089513.
15. Shrestha A., Xing L. D., Coit D. W. An efficient multistate multi-valued decision diagram-based approach for multistate system sensitivity analysis. IEEE Transactions on Reliability 2010; 59(3): 581-592, https://doi.org/10.1109/TR.2010.2055922.
16. Shrestha A., Xing L. D., Coit D. W. Multi-state component importance analysis using multi-state multi-valued decision diagrams. Proceedings of the 8th International Conference on Reliability, Maintainability and Safety 2009: 99-103, https://doi.org/10.1109/icrms.2009.5270231.
17. Shrestha A., Xing L. D., Dai Y. Decision diagram based methods and complexity analysis for multi-state systems. IEEE Transactions on Reliability 2010; 59(1): 145-161, https://doi.org/10.1109/TR.2009.2034946.
18. Tang Z., Dugan J. B. BDD-based reliability analysis of phased-mission systems with multimode failures. IEEE Transactions on Reliability 2006; 56(2): 350-360, https://doi.org/10.1109/TR.2006.874941.
19. Wang C. N., Xing L. D. Competing failure analysis in phased-mission systems with functional dependence in one of phases. Reliability Engineering & System Safety 2012; 108: 90-99, https://doi.org/10.1016/j.ress.2012.07.004.
20. Xing L. D. Reliability evaluation of phased-mission systems with imperfect fault coverage and common-cause failures. IEEE Transactions on Reliability 2007; 56(1): 58-68, https://doi.org/10.1109/TR.2006.890900.
21. Xing L. D. BDD-based reliability evaluation of phased-mission systems with internal/external common-cause failures. Reliability Engineering & System Safety 2013; 112: 145-153, https://doi.org/10.1016/j.ress.2012.12.003.
22. Xing L. D. and Dai Y. A new decision diagram based method for efficientanalysis on multi-state systems. IEEE Transactions Dependable &Secure Computing 2009; 6(3): 161-174, https://doi.org/10.1109/TDSC.2007.70244.
23. Xing L. D., Amari S. Reliability of k-out-of-n systems with phased-mission requirements and imperfect fault coverage. Reliability Engineering & System Safety 2012; 103: 45-50, https://doi.org/10.1016/j.ress.2012.03.018.
24. Zang X., Sun H., Trivedi K. S. A BDD-based algorithm for reliability analysis of phased-mission systems. IEEE Transactions on Reliability 1999; 48(1): 50-60, https://doi.org/10.1109/24.765927.
25. Zhang T., Bai G. H. Success probability model of phased mission systemswith limited spares. Eksploatacja i Niezawodnosc - Maintenance and Relaibility 2012; 14(1): 24-32.

Shuai ZHANG
Shudong SUN
Shubin SI
School of Mechantronics
Northwestern Polytechnical University, Xi'an, China
554 Mail Box, 127 West Youyi Road Xi'an Shaanxi, 710072, P.R. China

Peng WANG
3988 Lian Hua Nan Lu, Shanghai, China, Zip Code 200241

E-mails: zhangshuai5000@nwpu.edu.cn, sdsun@nwpu.edu.cn, sisb@nwpu.edu.cn, pengwang2005@yahoo.com

INFORMATION FOR AUTHORS

Eksploatacja i Niezawodnosc – Maintenance and Reliability – the journal of the Polish Maintenance Society, under the scientific supervision of the Polish Academy of Sciences (Branch in Lublin), published four times a year.

The scope of the Quarterly

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* publishes articles containing original results of experimental research on the durability and reliability of technical objects. We also accept papers presenting theoretical analyses supported by physical interpretation of causes or ones that have been verified empirically. *Eksploatacja i Niezawodność – Maintenance and Reliability* also publishes articles on innovative modeling approaches and research methods regarding the durability and reliability of objects.

The following research areas are particularly relevant to the journal:

1. degradation processes of mechanical and biomechanical systems,
2. diagnosis and prognosis of operational malfunctions and failures.
3. analysis of failure risk/wear,
4. reliability-and-environmental-safety engineering in the design, manufacturing and maintenance of objects,
5. management and rationalization of object maintenance,
6. risk management in the processes of operation and maintenance,
7. the human factor and human reliability in operation and maintenance systems.

Terms and Conditions of Publication

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* publishes only original papers written in English or in Polish with an English translation. Translation into English is done by the Authors after they have received information from the Editorial Office about the outcome of the review process and have introduced the necessary modifications in accordance with the suggestions of the referees! Acceptance of papers for publication is based on two independent reviews commissioned by the Editor.

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* proceeds entirely online at submission.ein.org.pl

Technical requirements

- After receiving positive reviews and after acceptance of the paper for publication, the text must be submitted in a Microsoft Word document format.
- Drawings and photos should be additionally submitted in the form of high resolution separate graphical files in the TIFF, SVG, AI or JPG formats.
- A manuscript should include: names of authors, title, abstract, and key words that should complement the title and abstract (in Polish and in English), the text in Polish and in English with a clear division into sections (please, do not divide words in the text); tables, drawings, graphs, and photos included in the text should have descriptive two-language captions, if this can be avoided, no formulae and symbols should be inserted into text paragraphs by means of a formula editor; references (written in accordance with the required reference format); author data – first names and surnames along with scientific titles, affiliation, address, phone number, fax, and e-mail address.

The Editor reserves the right to abridge and adjust the manuscripts. All submissions should be accompanied by a submission form.

Detailed instructions to Authors, including evaluation criteria can be found on the journal's website: www.ein.org.pl

Editor contact info

Editorial Office of „Eksploatacja i Niezawodnosc - Maintenance and Reliability"
Nadbystrzycka 36, 20-618 Lublin, Poland
e-mail: office@ein.org.pl

INFORMATION FOR SUBSCRIBERS

Fees

Yearly subscription fee (four issues) is 100 zloty and includes delivery costs. Subscribers receive any additional special issues published during their year of subscription free of charge.

Orders

Subscription orders along with authorization to issue a VAT invoice without receiver's signature should be sent to the Editor's address.

Note

In accordance with the requirements of citation databases, proper citation of publications appearing in our Quarterly should include the full name of the journal in Polish and English without Polish diacritical marks, i.e.,

Eksploatacja i Niezawodnosc – Maintenance and Reliability.

Wydawca:
Polskie Naukowo Techniczne
Towarzystwo Eksploatacyjne
Warszawa

Publisher:
Polish Maintenance Society
Warsaw

Członek:
Europejskiej Federacji
Narodowych Towarzystw
Eksploatacyjnych

Member of:
European Federation of National
Maintenance Societies

Patronat naukowy:
Polska Akademia Nauk
Oddział Lublin

Scientific Supervision:
Polish Academy of Sciences
Branch in Lublin