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 IF is 0.333

Editorial staff:	Dariusz Mazurkiewicz, PhD, DSc (Eng), Associate Professor (Editor-in-Chief, Secretary of the Scientific Board) Tomasz Klepka, PhD, DSc (Eng) (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 20-618 Lublin, ul. Nadbystrzycka 36, Poland e-mail: office@ein.org.pl http://www.ein.org.pl/
Circulation:	550 copies

Science and Technology

GAO W, ZHANG Z, ZHOU Y, WANG R, JIANG L. **Optimal combinatorial replacement policy under a given maintenance interval for the combined governor in diesel locomotives**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 89–98.

Combined governor is one of key components in diesel locomotives, as a subcomponent it must meet the existing maintenance periodic of the diesel locomotive, while it is passively replaced or maintained in a midway because of over/under-maintenance in practice. In this paper, four reliability models of a sequential PM cycle are developed using three years of maintenance data of combined governors in one Chinese Railway Bureau to determine its reliability distribution, in which some zero-failure data and censor data are used. Meanwhile, a novel combinatorial replacement (CR) policy is proposed to optimize its preventive maintenance (PM), in which a component is replaced several times using a preventively maintained one in a given operational interval. After that, necessary optimizations are introduced based on the determined reliability models and the maintenance interval given by the PM criterion of diesel locomotives (23000km ~25000km), and then the genetic algorithm is also used to solve the constraint optimization function. Results show that the proposed CR policy is the best policy among the existing policy and the general (T,δ) policy, and other results can be viewed as an optional policy when spare components are limited.

GAO W, ZHANG Z, ZHOU Y, WANG R, JIANG L. **Optymalna strategia wymiany kombinatorycznej dla danej częstotliwości przeglądów regulacji mieszanej w lokomotywach spalinowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 89–98.

Regulacja mieszana jest jednym z kluczowych elementów w lokomotywach spalinowych, i jako taka musi być ujęta w istniejącym systemie konserwacji okresowej lokomotyw spalinowych. W praktyce jednak, podlega ona biernej wymianie lub przedwczesnej konserwacji z powodu niewystarczających lub nadmiernych praktyk utrzymania w ruchu. W niniejszym artykule opracowano cztery modele niezawodnościowe sekwencyjnego cyklu konserwacji zapobiegawczej z wykorzystaniem danych z trzech lat konserwacji mieszanej regulacji jednego Biura Kolei Chińskich w celu ustalenia ich rozkładu niezawodności, przy użyciu wybranych danych nt. nieuszkadzalności oraz danych cenzurowanych. W celu optymalizacji konserwacji zapobiegawczej (PM), zaproponowano nową strategię wymiany kombinatorycznej (combinatorial replacement - CR), w której element składowy zostaje kilkakrotnie zastąpiony innym elementem uprzednio poddanym konserwacji zapobiegawczej w danym okresie eksploatacyjnym. Następnie, wprowadzono konieczne optymalizacje na podstawie opracowanych modeli niezawodnościowych oraz częstotliwości przeglądów podanej w kryteriach konserwacji zapobiegawczej lokomotyw spalinowych (23000 km ~ 25000 km). W dalszej kolejności, wykorzystano algorytm genetyczny do rozwiązania funkcji optymalizacji ograniczeń. Wyniki pokazują, że proponowana strategia CR jest najlepsza spośród istniejących strategii i ogólnych strategii (T, δ); inne wyniki można traktować jako strategię opcjonalną w sytuacji gdy dostępność elementów zamiennych jest ograniczona.

FUCHS P, ZAJICEK J. **Safety integrity level (SIL) versus full quantitative risk value**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 99–105.

Safety management of technical equipment is not possible without risk assessment. Therefore, many standards are available for risk assessment, e.g. ISO 13824:2009 General principles on risk assessment of systems involving structures or ISO/IEC 31010:2009 Risk management – Risk Assessment Techniques. In different industrial sectors risk assessment is fundamental step to determine required safety integrity level (SIL), eventually performance level (PL), which guarantees risk linked to some equipment on acceptable level. Standards applied for risk management based on SIL in different industrial sectors differ in methods used for risk evaluation and SIL determination. IEC 61508-5 accepts the use of qualitative, semi-quantitative or quantitative approach for risk evaluation and SIL determination. The standard uses hazardous event severity matrix as an example of qualitative approach for SIL determination, the standard furthermore uses layer of protection analysis (LOPA) as an example of semi-quantitative approach. The standard also uses Risk graph method as an example of both qualitative and semi-quantitative approach. IEC 62061 only presents one semi-quantitative approach for risk evaluation and SIL determination based on combination of probability and severity of consequences. This approach is different from the approach presented in IEC 61508-5. Similarly ISO 13849-1 recommends the use of qualitative method combining probability and severity of consequences for risk evaluation and PL determination, however, distinctly from IEC 61508-5 as well as IEC 62061. All these standards evaluate risk in the first step and in the second step they set safety systems reliability requirements, which should lower risk onto an acceptable level. The elemental question is, how exactly these standards evaluate risk in their methods. Another question is what acceptable level of risk is implicitly hidden in their requirements for choice of SIL and PL. This paper addresses these questions.

FUCHS P, ZAJICEK J. **Nienaruszalność bezpieczeństwa a wartość ryzyka**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 99–105.

Zarządzanie bezpieczeństwem urządzeń technicznych nie jest możliwe bez oceny ryzyka. Dlatego też istnieje wiele norm związanych z oceną ryzyka, np. ISO 13824:2009 Ogólne zasady dotyczące oceny ryzyka w systemach obejmujących konstrukcje lub ISO/IEC 31010:2009 Zarządzanie ryzykiem - Techniki oceny ryzyka. W różnych gałęziach przemysłu ocena ryzyka jest podstawowym krokiem na drodze do określenia wymaganego poziomu nienaruszalności bezpieczeństwa (SIL), oraz ewentualnie poziomu wydajności (PL), który gwarantuje, że ryzyko w odniesieniu do niektórych urządzeń pozostanie na akceptowalnym poziomie. Normy stosowane w zakresie zarządzania ryzykiem w oparciu o SIL w różnych gałęziach przemysłu różnią się jeśli chodzi o metody stosowane do oceny ryzyka i określenia SIL. IEC 61508-5 akceptuje zastosowanie jakościowego, pół-ilościowego lub ilościowego podejścia do oceny ryzyka oraz określenia SIL. Norma ta wykorzystuje macierz ciężkości zdarzeń niebezpiecznych jako przykład podejścia jakościowego do określenia SIL; ponadto, norma wykorzystuje analizę warstw zabezpieczeń (LOPA) jako przykład podejścia półilościowego. Norma wykorzystuje również metodę wykresu ryzyka jako przykład podejścia zarówno jakościowego jak i półilościowego. IEC 62061 prezentuje jedno pół-ilościowe podejście do oceny ryzyka i określenia SIL łącząc prawdopodobieństwo i ciężkość następstw. To podejście różni się od metody stosowanej w IEC 61508-5. Podobnie ISO 13849-1 zaleca stosowanie metody jakościowej łączącej prawdopodobieństwo i ciężkość następstw dla oceny ryzyka i określenia PL, jednak w sposób odmienny od IEC 61508-5 oraz IEC 62061. Wszystkie powyższe normy dokonują oceny ryzyka w pierwszym etapie zaś w drugim etapie ustalają one wymagania odnośnie niezawodności systemów bezpieczeństwa, które powinny obniżyć ryzyko do akceptowalnego poziomu. Podstawowym pytaniem jest jak dokładnie powyższe normy dokonują oceny ryzyka przy użyciu swoich metod. Inną kwestią jest to, jaki dopuszczalny poziom ryzyka jest domyślnie ukryty w ramach ich wymagań dotyczących wyboru SIL i PL. Niniejszy artykuł odnosi się do powyższych zagadnień.

DĘBSKI H. **Experimental investigation of post-buckling behavior of composite column with top-hat cross-section**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 106–110.

The object of this study is a thin-walled beam made of carbon-epoxy composite with open cross-section. The material used was a composite of epoxy matrix reinforced with carbon fiber (system HexPly M12, Hexcel). The M12 system is used above all in aircraft structures. It exhibits high fatigue durability and good maintenance properties at relatively low specific gravity. The research was lead as the FEM numerical analyses and experimental tests in buckling and post-buckling state, as well. In the conducted research in order to evaluate the effort ratio of the composite the Tsai-Wu tensor criterion was exploited. The numerical tool used was the ABAQUS software.

DĘBSKI H. **Eksperymentalno-numeryczne badania pokrytycznych zachowań kompozytowych kolumn o przekroju omegowym**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 106–110.

Przedmiotem badań jest cienkościenna belka wykonana z kompozytu węglowo-epoksydowego o przekroju otwartym. Zastosowanym materiałem był kompozyt o osnowie żywicy epoksydowej wzmacniany włóknami węglowymi systemu HexPly M12 (Hexcel). System M12 wykorzystywany jest w przede wszystkim w strukturach lotniczych i charakteryzuje się wysoką trwałością zmęczeniową oraz dobrymi właściwościami eksploatacyjnymi, przy stosunkowo niskim ciężarze własnym. Badania prowadzono w zakresie obliczeń numerycznych z wykorzystaniem MES oraz badań eksperymentalnych w stanie krytycznym i pokrytycznym. W prowadzonych badaniach do oceny stopnia wytężenia kompozytu wykorzystano kryterium tensorowe Tsai-Wu. Zastosowanym narzędziem numerycznym był program ABAQUS.

KAMARUZZAMAN S-N, MYEDA NE, PITT M. **Performance levels of high-rise private office buildings maintenance management in Malaysia**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 111–116.

Maintenance management is an essential aspect in determining the performance and quality of properties such as office buildings. The fundamental issues related to techniques and approaches concerned are generally taken lightly by the practitioners

KAMARUZZAMAN S-N, MYEDA NE, PITT M. **Poziomy wydajności zarządzania utrzymaniem ruchu w prywatnych wielokondygnacyjnych budynkach biurowych w Malezji**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 111–116.

Zarządzanie utrzymaniem ruchu jest ważnym aspektem przy określaniu wydajności i jakości mienia, takiego jak na przykład budynki biurowe. Podstawowe zagadnienia związane z odpowiednimi technikami i metodami są na ogół traktowane dość pochopnie przez which lead to inefficiency of maintenance management practice in the market today. The paper aims to determine the current standard and performance of maintenance management system by applying the study to high-rise private office buildings. These objectives are to be achieved by evaluating and analysing perceptions of the end users from five (5) high-rise office buildings in Klang Valley by using a mixed method combination of both quantitative and qualitative methods. Research findings signify that the performance of high-rise office buildings are generally rated as average by the end users and results from interviews with the maintenance managers denote the detailed systems applied. There is a positive relationship between the maintenance management systems and performance of maintenance management. This paper provides an important research which uncovered the scenario in the industry and the key perceptions by the building end users. This research is anticipated to be significantly beneficial and can be further used as a piece of information specifying on high rise private office buildings.

praktyków, co może prowadzić do powszechnej nieefektywności zarządzania utrzymaniem ruchu. Celem artykułu jest określenie bieżącego poziomu i wydajności systemu zarządzania utrzymaniem ruchu poprzez badania, których przedmiotem są wielokondygnacyjne prywatne budynki biurowe. Cele te zrealizowano poprzez ocenę i analizę obserwacji użytkowników końcowych z pięciu wielokondygnacyjnych budynków biurowych w Dolinie Klang przy użyciu metody mieszanej łączącej podejście ilościowe i jakościowe. Wyniki badań pokazują, iż wydajność wielokondygnacyjnych budynków biurowych jest ogólnie oceniana przez użytkowników końcowych jako przeciętna, podczas gdy rezultaty wywiadów ze specjalistami utrzymania ruchu dostarczyły informacji nt. stosowanych systemów. Istnieje pozytywny związek między systemami zarządzania utrzymaniem ruchu oraz jego wydajnością. Niniejszy artykuł przedstawia ważne badania, które prezentują praktyki stosowane w przemyśle jak również kluczowe spostrzeżenia użytkowników końcowych budynków. Badania mogą mieć korzystne przełożenie; mogą także znaleźć dalsze zastosowanie jako źródło szczegółowych informacji nt. wielokondygnacyjnych prywatnych budynków biurowych.

LI Y-F, MI J, HUANG H-Z, XIAO N-C, ZHU S-P. **System reliability modeling and assessmentfor solar array drive assembly based on bayesian networks**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 117–122.

Along with the increase of complexity in engineering systems, there exist many dynamic characteristics within the system failure process, such as sequence dependency, functional dependency and spares. Markov-based dynamic fault trees can figure out the modeling of systems with these characteristics. However, when confronted with the issue of state space explosion resulted from the growth of system complexity, the Markov-based approach is no longer efficient. In this paper, we combine the Bayesian networks with the dynamic fault trees to model the reliability of such types of systems. The inference technique of Bayesian network is utilized for reliability assessment and fault probability estimation. The solar array drive assembly is used to demonstrate the effectiveness of this method.

LI Y-F, MI J, HUANG H-Z, XIAO N-C, ZHU S-P. **Modelowanie i ocena niezawodności systemu w oparciu o sieci bayesowskie na przykładzie układu napędu paneli słonecznych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 117–122.

Wraz ze wzrostem złożoności w systemach technicznych, pojawia się wiele charakterystyk dynamicznych w ramach procesu awarii systemu, takich jak zależność sekwencyjna, zależność funkcjonalna czy zabezpieczające elementy zapasowe. Oparte na koncepcjach Markowa dynamiczne drzewa uszkodzeń mogą posłużyć do modelowania systemów z powyższymi charakterystykami. Jednak w konfrontacji z problemem eksplozji stanów wynikającym ze wzrostu złożoności systemu, podejście oparte na teoriach Markowa nie jest już skuteczne. W niniejszej pracy łączymy sieci bayesowskie z dynamicznymi drzewami uszkodzeń w celu modelowania niezawodności tego typu systemów. Technikę wnioskowania sieci bayesowskiej wykorzystano do oceny niezawodności i prawdopodobieństwa wystąpienia uszkodzenia. Skuteczność niniejszej metody wykazano na przykładzie układu napędu paneli słonecznych.

TARASIUK P, MOLSKI KL, SZYMANIUK A. **Fatigue designing of welded agricultural wheels**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 123–128.

Each introduction of new or modified products into production requires conducting a series of calculations and experiments which would verify the desired quality. In case of agricultural wheels it is necessary to estimate their fatigue life and durability at a possibly low cost. This paper presents a method of fatigue design of agricultural wheels on the example of a welded wheel type 9.00x15.3. A previously designed numerical model FEM was used, which enables to identify the potential weak spots of the construction which could determine its total durability and allows to assess the critical parameters. Based on the results of model studies, construction changes were introduced, which were experimentally verified on a fatigue machine during durability testing for radial forces.

TARASIUK P, MOLSKI KL, SZYMANIUK A. **Projektowanie trwałościowe spawanych kół pojazdów rolniczych**. Eksploatacja i Niezawodnosc – Mainte nance and Reliability 2013; 15 (2): 123–128.

Wprowadzenie do produkcji nowych lub zmienionych pod względem konstrukcyjnym wyrobów wymaga wykonania obliczeń sprawdzających oraz doświadczeń potwierdzających uzyskanie ich założonej jakości. W przypadku kół pojazdów rolniczych konieczne jest wcześniejsze oszacowanie ich wytrzymałości i trwałości zmęczeniowej możliwie niskim kosztem. W pracy przestawiono metodę projektowania trwałościowego kół pojazdów wolnobieżnych na przykładzie spawanego koła typu 9.00x15.3. Wykorzystano opracowany wcześniej model numeryczny MES, który umożliwia identyfikację potencjalnie najsłabszych miejsc konstrukcji decydujących o jej trwałości oraz pozwala określić wartości parametrów krytycznych. Na podstawie wyników badań modelowych wprowadzono zmiany konstrukcyjne, które zweryfikowano doświadczalnie na stanowisku badawczym w testach trwałościowych na obciążenia promieniowe.

GRONOSTAJSKI Z, HAWRYLUK M, KRAWCZYK J, MARCINIAK M. **Numerical modelling of the thermal fatigue of steel WCLV used for hot forging dies**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 129–133.

The paper presents an analysis of a numerical simulation of the low-cycle thermal fatigue of steel WCLV (X40CrMoV511) used in hot forging. As part of experimental studies a special test rig based on the rotating disc method was built and tests were carried out. Their resultsshowed that the method can be used to reproducethe thermal fatigue conditions prevailing in the industrial forging process. For the given experimental conditions the instant when fatigue cracks appear was determined. A numerical model was built and the obtained finite element analysis results were compared with the laboratory test results in order to determine the amplitude of plastic strains at which a crack appears. As part of further research in the future the Coffin-Manson low-cycle fatigue model will be verified for other conditions and a low-cycle fatigue curve for steel WCLV will be determined.

GRONOSTAJSKI Z, HAWRYLUK M, KRAWCZYK J, MARCINIAK M. **Modelowanie numeryczne zmęczenia cieplnego stali WCLV stosowanej na matryce w procesie kucia na gorąco**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 129–133.

W pracy przedstawiono analizę symulacji numerycznej niskocyklowego zmęczenia cieplnego stali WCLV stosowanej podczas kucia na gorąco. W ramach badań doświadczalnych zostało zbudowane specjalne stanowisko bazujące na metodzie "wirującego krążka" [13], przeprowadzone zostały próby, które potwierdziły możliwość stosowania tej metody do odwzorowania warunków zmęczenia cieplnego panujących w przemysłowym procesie kucia. Dla danych warunków eksperymentu określono moment pojawienia się pęknięć zmęczeniowych. Następnie zbudowano model numeryczny, po czym porównano uzyskane wyniki z MES i prób laboratoryjnych w celu określenia amplitudy odkształceń plastycznych, przy których pojawia się pęknięcie. Dalsze prace pozwolą w przyszłości na weryfikację niskocyklowego modelu zmęczenia Coffina-Mansona dla innych warunków oraz pozwolą stworzyć krzywą niskocyklowego zmęczenia stali WClV.

PARCZEWSKI K. **Effect of tyre inflation preassure on the vehicle dynamics during braking manouvre**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 134–139.

The paper presents the problem of reducing the impact of inflation pressure on the tires, and the weight distribution of the vehicle on the road. The results presented in this publication are based on the research bench and passenger vehicle equipped with anti-lock braking system. Bench testing was conducted stiffness of tires and road tests, performing maneuvers based on ISO standards, braking in the straight patch of road and of the twisting road.

PARCZEWSKI K. **Wpływ ciśnienia w ogumieniu na dynamikę ruchu pojazdu podczas manewru hamowania**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 134–139.

W publikacji przedstawiono zagadnienie wpływu obniżenia ciśnienia w oponie na charakterystykę opon, rozkład nacisków i zachowanie się pojazdu na drodze. Wyniki prezentowane w publikacji oparto na badaniach stanowiskowych i drogowych samochodu osobowego wyposażonego w układ zapobiegający blokowaniu kół. Przeprowadzono badania stanowiskowe sztywności opon oraz badania drogowe, polegające na wykonywaniu manewrów opartych na normach ISO: hamowania na prostoliniowym odcinku oraz na łuku drogi.

SKOWRONEK K, WOŹNIAK A. **FFT-PCA-LDA classifier in A.C. Generator diagnostics**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 140–146.

The methods of A.C. generator diagnostics are discussed. The need of developing new methods is justified. A new classification method is presented that is used for diagnostics of A.C. generator damages. Features of the method are specified. Functioning of the method is analyzed based on examination of damages of A.C. generator diodes. The method is compared with other methods of electric machine diagnostics used in practice.

SKOWRONEK K, WOŹNIAK A. **Klasyfikator FFT-PCA-LDA w diagnostyce alternatora**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 140–146.

Omówiono metody diagnostyki alternatorów. Uzasadniono konieczność konstrukcji nowych metod. Zaprezentowano nową metodę klasyfikacyjną wykorzystaną do diagnostyki uszkodzeń alternatora. Przedstawiono cechy metody. Działanie metody przeanalizowano na podstawie badania uszkodzeń diod alternatora. Metodę porównano z metodami diagnostyki maszyn elektrycznych stosowanymi w praktyce.

YANG Z-J, CHEN C-H, CHEN F, HAO Q-B, XU B-B. **Reliability analysis of machining center based on the field data**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 147–155.

Machining center is the complex machinery, with high level automation and complicated structures, so there are lots of failures. When a random failure occurs, the failed machining center stops and causes a production line or even the whole workshop to stop functioning. The frequent failure leads to the low levels of reliability and production rate. In order to help users and manufacturers optimize maintenance policy to improve the reliability for machining center, this paper presents descriptive statistics of the failure data and develops the failure trend using power-law process, simultaneously establishes the routine inspection and regular inspection as well as the sequential preventive maintenance under maintenance cost constraints. The proposed model could be a useful tool to assess the current conditions, predict reliability and optimize the machining center maintenance policy.

YANG Z-J, CHEN C-H, CHEN F, HAO Q-B, XU B-B. **Analiza niezawodnościowa centrum obróbkowego w oparciu o dane terenowe**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 147–155.

Centrum obróbkowe to skomplikowany mechanizm o wysokim poziomie automatyzacji oraz złożonej konstrukcji, w związku z czym ulega licznym uszkodzeniom. Przy wystąpieniu przypadkowej awarii, uszkodzone centrum obróbkowe przestaje działać i powoduje zatrzymanie linii produkcyjnej a nawet całego oddziału produkcyjnego. Częste awarie obniżają poziom niezawodności oraz tempo produkcji. Aby pomóc użytkownikom i producentom zoptymalizować politykę utrzymania ruchu w celu poprawy niezawodności centrów obróbkowych, w niniejszym artykule przedstawiono statystyki opisowe dotyczące danych o uszkodzeniach i opracowano trend uszkodzeń w oparciu o proces spełniający prawo potęgowe. Jednocześnie ustalono zasady rutynowej inspekcji i okresowych przeglądów, jak również sekwencyjnej obsługi zapobiegawczej przy ograniczonych wydatkach na utrzymanie ruchu. Proponowany model może być użytecznym narzędziem dla potrzeb oceny aktualnych warunków oraz przewidywania niezawodności w celu optymalizacji polityki utrzymania ruchu centrum obróbkowego.

GIRTLER J, ŚLĘZAK M. **Four-state stochastic model of changes in the reliability states of a motor vehicle**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 156–160.

The properties of semi-Markov processes have been generally characterized and the applicability of the theory of such processes to the determining of the reliability of motor cars and other road vehicles has been explained. A formal description of the process of changes in the motor vehicle technical states considered as reliability states and a model of this process in the form of a one-dimensional stochastic process have been presented. The values of this process are the technical states of the motor vehicle in question that have significant practical importance. A four-state set of states interpreted as follows has been adopted: full (complete) serviceability, partial (incomplete) serviceability,task-limiting serviceability,and complete (total) unserviceability. Based on the initial distribution adopted and the functional matrix worked out, the boundary distribution of the process of changes in the technical (reliability) states of the motor vehicle has been defined. The probability of the vehicle being fully serviceable has been considered a measure of the vehicle reliability for a long period of vehicle operation. A possibility of defining the vehicle reliability in the form of a probability that a task would also be fulfilled by the vehicle being partially serviceable has also been indicated.

GIRTLER J, ŚLĘZAK M. **Model stochastyczny czterostanowy zmian stanów niezawodnościowych samochodu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 156–160.

W artykule scharakteryzowano ogólnie własności procesów semimarkowskich i uzasadniono możliwości ich zastosowania do określenia niezawodności samochodów i innych pojazdów drogowych. Przedstawiono formalny opis procesu zmian stanów technicznych samochodów uznanych za stany niezawodnościowe oraz model tego procesu w postaci jednowymiarowego procesu stochastycznego. Wartościami tego procesu są występujące w czasie eksploatacji stany techniczne samochodów, mające istotne znaczenie praktyczne. Przyjęto czterostanowy zbiór stanów o następującej interpretacji: stan zdatności pełnej (całkowitej), stan zdatności częściowej (niepełnej, niecałkowitej), stan niepełnej zdatności zadaniowej i stan niezdatności pełnej (całkowitej). Na podstawie przyjętego rozkładu początkowego i opracowanej macierzy funkcyjnej został określony rozkład graniczny procesu zmian stanów technicznych (niezawodnościowych) samochodu. Prawdopodobieństwo istnienia stanu zdatności pełnej (całkowitej) samochodu zostało uznane za miarę jego niezawodności w długim okresie czasu eksploatacji. Wskazano też na możliwość określenia niezawodności samochodu w formie prawdopodobieństwa, w którym uwzględniony został przypadek wykonania zadania przez samochód także wtedy, gdy znajduje się on w stanie zdatności częściowej.

AMANI N, ALI NM, MOHAMMED AH, SAMAT RA. **Maintenance and management of wastewater system components using the condition index system, prediction processand costs estimation**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 161–168.

Component maintenance management of public building is complex and dynamic as the execution of the engineering management system is subjected to sensitive staff and users' requirements and high expectation of the top management for supporting the business. This paper presents the practices and survey need for maintaining the facilities systems in the building. The purpose of this study is maintenance time optimization of building component using the USACERL condition index (CI) system. To achieve this objective, cast iron pipe within wastewater plumbing system is surveyed using the financial analysis for implementation of optimal maintenance time based on limited cost. The findings show how a best time approach to plumbing system maintenance can assist the owner for decision making in component maintenance time based on existing cost.

AMANI N, ALI NM, MOHAMMED AH, SAMAT RA. **Konserwacja i zarządzanie systemem kanalizacji ściekowej za pomocą systemu wskaźnika stanu, procesu przewidywania i szacowania kosztów**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 161–168.

Zarządzanie konserwacją instalacji budynku publicznego jest złożone i dynamiczne z uwagi na fakt, iż realizacja systemu zarządzania technicznego poddana jest zarówno wymaganiom personelu i użytkowników końcowych jak i oczekiwaniom kierownictwa w zakresie wsparcia rozwoju przedsiębiorstwa. W niniejszym artykule przedstawiono praktykę oraz badania dotyczące potrzeb wiążących się z konserwacją instalacji w budynkach. Celem tego opracowania jest optymalizacja czasu konserwacji tej części budynku za pomocą systemu wskaźnika stanu (condition index - CI) USACERL. Aby osiągnąć ten cel, przeanalizowano za pomocą analizy finansowej system kanalizacji ściekowej oparty na rurach żeliwnych pod kątem przyjęcia optymalnego czasu obsługi w oparciu o ograniczone koszty. Wyniki pokazują jak optymalne podejście czasowe do konserwacji systemu hydraulicznego może pomóc właścicielowi w procesie podejmowania decyzji w aspekcie czasu konserwacji na podstawie rzeczywistych kosztów.

MI J, LI Y, HUANG H-Z, LIU Y, ZHANG X-L. **Reliability analysis of multi-state system with common cause failure based on bayesian networks**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 169–175.

Taking account of the influence of common cause failure (CCF) to system reliability and the widespread presence of multi-state system (MSS) in engineering practices, a method for reliability modeling and assessment of a multi-state system with common cause failure is proposed by taking the advantage of graphic representation and uncertainty reasoning of Bayesian Network (BN). The model is applied to a two-axis positioning

MI J, LI Y, HUANG H-Z, LIU Y, ZHANG X-L. **Analiza niezawodności systemu wielostanowego z uszkodzeniem spowodowanym wspólną przyczyną w oparciu o sieci bayesowskie**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 169–175.

Uwzględniając wpływ uszkodzeń spowodowanych wspólną przyczyną (CCF) na niezawodność systemów oraz powszechne występowanie w praktyce inżynierskiej systemów wielostanowych (MSS), zaproponowano metodę modelowania i oceny niezawodności systemu wielostanowego z uszkodzeniem spowodowanym wspólną przyczyną, która wykorzystuje reprezentację graficzną sieci Bayesa (BN) i oparte na nich wnioskowanie mechanism transmission system to demonstrate its effectiveness and capability for directly calculating the system reliability on the basis of multi-state probabilities of components. Firstly, the reliability block diagram is built according to the hierarchy of structure and function of multi-state system. Then, the traditional Bayesian Networks model of the transmission system is constructed based on the reliability block diagram, failure logic between components and the failure probability distribution of them. In this paper, the β-factor model is used to analyze the CCF of the transmission system, and a new Bayesian network combining with CCF is established following by the implementation of reliability analysis. Finally, the comparison between the proposed method and the one without considering CCF is made to verify the efficiency and accuracy of the proposed method.

przybliżone. Model zastosowano do analizy układu przenoszenia napędu dwu-osiowego mechanizmu pozycjonowania. Zbadano w ten sposób skuteczność modelu oraz możliwość wykorzystania go do bezpośredniego obliczania niezawodności systemu na podstawie wielostanowych prawdopodobieństw elementów składowych. W pierwszej kolejności stworzono schemat blokowy niezawodności uwzględniający hierarchię struktury i funkcji badanego systemu wielostanowego. Następnie, w oparciu o schemat blokowy niezawodności, logikę uszkodzeń komponentów oraz rozkład prawdopodobieństwa uszkodzeń tych komponentów, skonstruowano tradycyjny model bayesowski układu przenoszenia napędu. W niniejszej pracy wykorzystano model współczynnika β do analizy CCF układu przenoszenia napędu oraz opracowano nową sieć Bayesa uwzględniającą CCF, po czym przeprowadzono na ich podstawie analizę niezawodności. Skuteczność i dokładność proponowanej metody sprawdzono poprzez porównanie jej z metodą nie wykorzystującą CCF.

KADYAN MS. **Reliability and profit analysis of a single-unit system with preventive maintenance subject to maximum operation time**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 176–181.

This paper deals with the profit analysis of a reliability model for a single-unit system in which unit fails completely either directly from normal mode or via partial failure. The partially failed operating unit is shutdown after a maximum operation time for preventive maintenance. There is a single server who attends the system immediately whenever needed to conduct preventive maintenance at partial failure stage and repair at completely failure stage of the unit. The unit works as new after preventive maintenance and repair. The switch devices are considered as perfect. All random variables are assumed as independent and uncorrelated. The distribution of failure times, maximum operation time, preventive maintenance time and repair time are taken as general. Various reliability characteristics of interest are evaluated by using semi-Markov process and regenerative point technique. The tabular represantation of mean time to system failure(MTSF), availability and profit with respect to maximum rate of operation time has also been shown for a particular case.

KADYAN MS. **Analiza niezawodności i zysku dla systemu jednoelementowego z konserwacją zapobiegawczą poddanego maksymalnemu czasowi pracy**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 176–181.

W niniejszej pracy przedstawiono analizę zysku modelu niezawodności dla systemu jednoelementowego, w którym element ulega całkowitemu uszkodzeniu bezpośrednio z trybu normalnego lub pośrednio na skutek częściowego uszkodzenia. Częściowo uszkodzona działająca jednostka jest wyłączana po upłynięciu maksymalnego czasu pracy w celu przeprowadzenia konserwacji zapobiegawczej. Pojedynczy serwer wspomaga bezzwłocznie system w momencie wystąpienia potrzeby przeprowadzenia konserwacji zapobiegawczej na etapie częściowego uszkodzenia oraz naprawy na etapie uszkodzenia całkowitego. Element działa jak nowy, po konserwacji zapobiegawczej i naprawie. Stan przełączników sieciowych uznaje się za doskonały. Wszystkie zmienne losowe traktowano jako niezależne i nieskorelowane. Rozkład czasów uszkodzeń, maksymalnego czasu pracy, czasu konserwacji zapobiegawczej i czasu naprawy przyjęto jako ogólne. Wybrane parametry niezawodnościowe oceniano za pomocą procesu semimarkowskiego i techniki odnowy RPT. Dla poszczególnych przykładów przedstawiono także tabelaryczne zestawienie średniego czasu do uszkodzenia systemu (MTSF), gotowości i zysku w odniesieniu do maksymalnego czasu pracy.

GRABOŚ A, BORYGA M. **Trajectory planning of end-effector with intermediate point**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 182–187.

The article presents the Polynomial Cross Method (PCM) for trajectory planning of an end-effector with an intermediate point. The PCM is applicable for designing robot end-effector motion, whose path is composed of two rectilinear segments. Acceleration profile on both segments was described by the 7th-degree polynomial. The study depicts an algorithm for the method and the research results presented as the runs of resultant velocity, acceleration and linear jerk of the stationary coordinate system.

GRABOŚ A, BORYGA M. **Planowanie trajektorii ruchu chwytaka z punktem pośrednim**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 182–187.

W pracy zaprezentowano metodę PCM (Polynomial Cross Method) do planowania trajektorii ruchu chwytaka z punktem pośrednim. PCM ma zastosowanie do planowania ruchu chwytaka, którego tor składa się z dwóch odcinków prostoliniowych. Profil przyspieszenia na obu odcinkach opisany został wielomianem siódmego stopnia. W pracy przedstawiono algorytm metody oraz wyniki w postaci przebiegów prędkości, przyspieszenia i udaru liniowego.

GUO Y-M, RAN C-B, LI X-L, MA J-Z, ZHANG L. **Weighted prediction method with multiple time series using multi-kernel least squares support vector regression**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 188–194.

Least squares support vector regression (LS-SVR) has been widely applied in time series prediction. Based on the case that one fault mode may be represented by multiple relevant time series, we utilize multiple time series to enrich the prediction information hiding in time series data, and use multi-kernel to fully map the information into high dimensional feature space, then a weighted time series prediction method with multi-kernel LS-SVR is proposed to attain better prediction performance in this paper. The main contributions of this method include three parts. Firstly, a simple approach is proposed to determine the combining weights of multiple basis kernels; Secondly, the internal correlative levels of multiple relevant time series are computed to present the different contributions of prediction results; Thirdly, we propose a new weight function to describe each data's different effect on the prediction accuracy. The experiment results indicate the effectiveness of the proposed method in both better prediction accuracy and less computation time. It maybe has more application value.

GUO Y-M, RAN C-B, LI X-L, MA J-Z, ZHANG L. **Metoda ważonej predykcji wielokrotnych szeregów czasowych z wykorzystaniem wielojądrowej regresji wektorów wspierających metodą najmniejszych kwadratów (LS-SVR)**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 188–194.

Regresja wektorów wspierających metodą najmniejszych kwadratów (LS-SVR) jest szeroko stosowana w predykcji szeregów czasowych. Opierając się na fakcie, że jeden rodzaj niezdatności może być reprezentowany przez wiele relewantnych szeregów czasowych, w niniejszej pracy wykorzystano wielokrotne szeregi czasowe do wzbogacenia informacji predykcyjnych ukrytych w szeregach czasowych oraz posłużono się metodą uczenia wielojądrowego (multi-kernel) w celu mapowania informacji do wysoko wymiarowej przestrzeni cech, a następnie zaproponowano metodę ważonej predykcji wielokrotnych szeregów czasowych z wykorzystaniem wielojądrowej regresji LS-SVR służącą osiągnięciu lepszej wydajności prognozowania.Metoda składa się z trzech głównych części. Po pierwsze, zaproponowano prosty sposób określania łącznej wagi wielu jąder podstawowych. Po drugie, obliczono wewnętrzne poziomy korelacyjne wielokrotnych szeregów czasowych w celu przedstawienia różnego udziału wyników prognozowania. Po trzecie, zaproponowano nową funkcję wagi do opisu różnego wpływu poszczególnych danych na trafność predykcji. Wyniki doświadczenia wskazują na skuteczność proponowanej metody zarówno jeśli chodzi o lepszą trafność predykcji jak i krótszy czas obliczeniowy. Proponowane rozwiązanie ma potencjalnie dużą wartość aplikacyjną.

GU H, ZHAO J, ZHANG X. **Hybrid methodology of degradation feature extraction for bearing prognostics**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 195–201.

Hybrid methodology of degradation feature extraction was presented which may enable prediction of remaining useful life of a product. In this methodology, firstly, the signal was de-noised by wavelet analysis. Then the autoregressive model was used to remove the discrete frequencies from de-noised signal. Further, the residual signal which mainly

GU H, ZHAO J, ZHANG X. **Metodyka hybrydowa ekstrakcji cech degradacji do zastosowań w prognozowaniu czasu życia łożysk**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 195–201.

Przedstawiono hybrydową metodę ekstrakcji cech degradacji, która umożliwia przewidywanie pozostałego okresu użytkowania produktu. W tej metodyce, sygnał został najpierw odfiltrowany z wykorzystaniem analizy falkowej. Następnie, za pomocą modelu autoregresyjnego usunięto z pozbawionego szumów sygnału częstotliwości

contained impulsive fault signal was enhanced by minimum entropy deconvolution filter. The kurtosis was extracted which was taken as the feature for prognostics. At last, the empirical mode decomposition was used to reduce fluctuation of feature value and to extract the trend content. A case study was presented to verify the effectiveness of the proposed method.

dyskretne. W dalszej kolejności, sygnał resztkowy, który zawierał głównie impulsowy sygnał uszkodzenia został wzmocniony z zastosowaniem filtru dekonwolucji minimum entropii. Obliczono kurtozę, którą przyjęto jako cechę w procesie prognozowania. Na koniec, zastosowano empiryczną dekompozycję sygnału (EMD) w celu zmniejszenia wahań wartości cechy oraz w celu ekstrakcji trendu. Studium przypadku demonstruje efektywność proponowanej metody.

PASKA J. **Chosen aspects of electric power system reliability optimization**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 202–208.

Reliability is one of the most important criteria, which must be taken into consideration during planning and operation phases of an electric power system, especially in present situation of the power sector. This paper considers the optimization of electric power system reliability. The formalization of description of electric power system reliability level optimization is done as well as its practical solving components are given: diagram of value based reliability approach and estimation of customer damage costs resulting from insufficient reliability level.

PASKA J. **Wybrane aspekty optymalizacji niezawodności systemu elektroenergetycznego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 202–208.

Niezawodność jest jednym z najważniejszych kryteriów, które należy uwzględniać, zarówno podczas planowania rozwoju, jak też eksploatacji systemu elektroenergetycznego, szczególnie w obecnej sytuacji elektroenergetyki. Artykuł dotyczy optymalizacji niezawodności systemu elektroenergetycznego. Przedstawiono formalny opis matematyczny zagadnienia optymalizacji poziomu niezawodności systemu elektroenergetycznego oraz pewne elementy jego rozwiązania: schemat podejścia wartościowania niezawodności oraz szacowanie kosztów strat odbiorców z tytułu niedostatecznego poziomu niezawodności.

This page intentionally left blank

Article citation info:
GAO W, ZHANG Z, ZHOU Y, WANG R, JIANG L. Optimal combinatorial replacement policy under a given maintenance interval for the combined governor in diesel locomotives. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 89–98.

Wenke GAO
Zhisheng ZHANG
Yifan ZHOU
Rujiao WANG
Lei JIANG

OPTIMAL COMBINATORIAL REPLACEMENT POLICY UNDER A GIVEN MAINTENANCE INTERVAL FOR THE COMBINED GOVERNOR IN DIESEL LOCOMOTIVES

OPTYMALNA STRATEGIA WYMIANY KOMBINATORYCZNEJ DLA DANEJ CZĘSTOTLIWOŚCI PRZEGLĄDÓW REGULACJI MIESZANEJ W LOKOMOTYWACH SPALINOWYCH

Combined governor is one of key components in diesel locomotives, as a subcomponent it must meet the existing maintenance periodic of the diesel locomotive, while it is passively replaced or maintained in a midway because of over/under-maintenance in practice. In this paper, four reliability models of a sequential PM cycle are developed using three years of maintenance data of combined governors in one Chinese Railway Bureau to determine its reliability distribution, in which some zero-failure data and censor data are used. Meanwhile, a novel combinatorial replacement (CR) policy is proposed to optimize its preventive maintenance (PM), in which a component is replaced several times using a preventively maintained one in a given operational interval. After that, necessary optimizations are introduced based on the determined reliability models and the maintenance interval given by the PM criterion of diesel locomotives (23000km ~25000km), and then the genetic algorithm is also used to solve the constraint optimization function. Results show that the proposed CR policy is the best policy among the existing policy and the general (T,δ) policy, and other results can be viewed as an optional policy when spare components are limited.

***Keywords**: Reliability models, mixed-Weibull distribution, combinatorial replacement policy, combined governor of diesel locomotives.*

Regulacja mieszana jest jednym z kluczowych elementów w lokomotywach spalinowych, i jako taka musi być ujęta w istniejącym systemie konserwacji okresowej lokomotyw spalinowych. W praktyce jednak, podlega ona biernej wymianie lub przedwczesnej konserwacji z powodu niewystarczających lub nadmiernych praktyk utrzymania w ruchu. W niniejszym artykule opracowano cztery modele niezawodnościowe sekwencyjnego cyklu konserwacji zapobiegawczej z wykorzystaniem danych z trzech lat konserwacji mieszanej regulacji jednego Biura Kolei Chińskich w celu ustalenia ich rozkładu niezawodności, przy użyciu wybranych danych nt. nieuszkadzalności oraz danych cenzurowanych. W celu optymalizacji konserwacji zapobiegawczej (PM), zaproponowano nową strategię wymiany kombinatorycznej (combinatorial replacement – CR), w której element składowy zostaje kilkakrotnie zastąpiony innym elementem uprzednio poddanym konserwacji zapobiegawczej w danym okresie eksploatacyjnym. Następnie, wprowadzono konieczne optymalizacje na podstawie opracowanych modeli niezawodnościowych oraz częstotliwości przeglądów podanej w kryteriach konserwacji zapobiegawczej lokomotyw spalinowych (23000 km ~ 25000 km). W dalszej kolejności, wykorzystano algorytm genetyczny do rozwiązania funkcji optymalizacji ograniczeń. Wyniki pokazują, że proponowana strategia CR jest najlepsza spośród istniejących strategii i ogólnych strategii (T, δ); inne wyniki można traktować jako strategię opcjonalną w sytuacji gdy dostępność elementów zamiennych jest ograniczona.

***Słowa kluczowe**: Modele niezawodności, rozkład Weibulla, strategia wymiany kombinatorycznej, lokomotywa spalinowa o mieszanej regulacji.*

1. Background

There are 11,000 diesel locomotives (about 40% all locomotives) in China and most of them are still on service at the main or the branch railway line. The combined governor is one of their key components and plays a significant role in the transmission and control system, which is a typical mechanic-electronic-hydraulic set, and is composed of the mechanical driving part and hydraulic operating control system and electron control system. Its main functions include: a) to control the rotating speed and power of the engine automatically; b) to modify the fuel supply based on the signal of driver-controller, the load and with zero-failure data. Especially, the Weibull distribution is widely used in reliability and lifetime analysis due to its versatility, such as discussed in [8, 10, 11, 20, 21]. Attention has been paid to zero-failure data since the paper by Martz & Waller [3, 14] was published, and yet it is a focus in development and less application in reality. Miller et al, [15] introduced formulae for estimating the probability of failure for software when testing reveals no errors. Jiang et al, [7] proposed a method to combine the initial guess of the reliability with the estimation from zero-failure data, to acquire a more reliable estimate for the Weibull distribution, while the method is on the basis of some additional information. Han [5] applied the hierarchical Bayesian and the "E-Bayes" methods to estimate production reliability by zero-failure testing data. This method can be used to deal with some maintenance observations when there is no failure occurring. In practice, due to the limitation of the maintenance decision, some key components are maintained under zero-failure state, while zero-failure data is valuable operational information and ignored on some occasion, the combined governor of diesel locomotives is a case in point. Meanwhile, most existing research on maintenance problem assumed that the reliability of components is a two-fold Weibull distribution or a known distribution. While it is difficult to fit the lifetime distribution using the two-fold Weibull model in some real cases. Therefore, four novel models are proposed applying the zero-failure data and the censored failure data in this paper.

For the research on maintenance policy, most research have been widely investigated since the repair model was presented in 1960 by Barlow and Hunter [16]. Wang[19] summarized the recently years PM policies and classified them into five classes: Age-dependent PM policy, Periodic PM policy, Failure limit policy, Sequential PM policy and Repair limit policy. Most of the existing research focused on optimal replacement policies with new components from the literatures review, only several literatures considered the replacement components with used one. Tango[18] proposed a extended block replacement policy with used items based on the assumption that if items fail in $[(k-1)T, kT-v)$, they are replaced by new items, and if in $[kT-v, kT)$, they are replaced by used items. Sheu et al[17] introduced a extended block replacement policy with shock models and used items, in which an operating system is preventively replaced by new ones at times iT (i=1,2,…) independently of its failure history. If system fails in ((i-1) T, iT-δ) it is either replaced by a new one or minimally repaired, and if in [iT-δ, iT) it is either replaced by a used one or minimally repaired. Zhao et al[23] considered three imperfect PM policies at time T, shock numbers N, and damage k of a used system, and expected cost rates were obtained by using the techniques of cumulative processes and reliability theory. While some existing literatures assume that the effective age of a used component is known that is difficulty to obtain in practice, meanwhile, due to the requirement of precise and reliability, once the component is used and take-down, a PM action must be made before next operation in some cases, such as the combined governor. Furthermore, most research regarded that the maintenance periodic is a fixed value, but it is an interval (23000 km ~ 25000 km) in diesel locomotives. Therefore, the existing maintenance models are unfit for combined governors.

3. Reliability Model of Combined Governors

As mention above, due to the importance of the combined governor in locomotives, it must be preventively maintained at every PM process of diesel locomotives. Consider that the PM action might change its reliability distribution, observations are classified into four groups based on the frequency of PMs: new combined governors (before the first PM, group1), after the first PM but before the second PM (group 2), and after the second PM but before the third PM (group 3) and after the third PM (group 4). The overhauling cycle of diesel locomotives is divided into four operational phases by three PM actions, and the overhauling cycle is performed at the end of the last operational phase. In additional, the lifetime of components in the diesel locomotive is running mileage, and the maintenance interval is also running mileage. Thus their units are kilometer (km) in this paper.

3.1. Model Assumption

Taking the above into consideration and following the two hypotheses given below.

1) The hazard rate is zero for a new combined governor or one that suffered multiple PMs at the beginning.
2) Malfunctions can be detected upon its occurrence and can be removed by a minimal repair at once which will restore the function of equipment without changing the hazard rate, and the time of a minimal repair and PM and replacement can be ignored.

3.2. Data Preprocessing

In this section, the real running observations of the combined governor deprived from DF4 diesel locomotives of one Railway Bureau (2009–2011) are analyzed. The statistics data include the censored data and the failure data labeled with "+", part of them can be seen as Table 1, from which it can be found that without failure occurs during the first operational phase, these observations can be viewed as zero-failure data. The rest observations include failure data and censored data in the second to the forth operational phase.

3.3. Distribution Fitting

1) Reliability distribution for the first operation phase

In the first operational phase, the observations are zero-failure data. The Hierarchical Bayesian and "E-Bayes" methods are employed to describe the distribution of zero-failure data. The latter is found to be an easier approach to the description [5], the process of which goes as follows:

a) Let (s_i, t_i) be the zero failure data, and pi the failure probability at time t_i.
b) Assume that the prior distribution of pi is Beta distribution with density function

$$\pi(p_i \mid a,b) = \frac{p_i^{\,a-1}(1-p_i)^{b-1}}{B(a,b)}$$

Where $0 < p_i < 1$, B(a,b) $= \int_0^1 t^{a-1}(1-t)^{b-1}dt$, and a > 0, b > 0, and both a and b are hyper parameters and separately below

Table 1. Part Maintenance records of the combined governor

group1	28868	21194	22673	20051	21768	22074	28075	24611
group2	22521	23852+	21031	24304	21457	24915+	23363	8154+
group3	22436	22565	20205	24913	20549+	24092	23388	21383
group4	25159+	28330	7025+	4869+	23572	22286	24195+	22870

P.S: "+"is the failure data.

the rotate speed of the engine; and c) to match the excitation current of generator. Due to its important role in diesel locomotives and the limitation of the structure, it must follow the existing maintenance periodic of the existing diesel locomotives. The current Railway Technical Management Criterion stresses that it must be disintegrated for tests of all its springs, fly hammers, pistons and motor as well as for renewing of the oil in every PM action to improve its control precision and reliability. This maintenance policy may cause over/under-maintenance, and leads to passive replacement/maintenance or accidents in midway in practice. Meanwhile, the maintenance cost and malfunctions still become outstanding problems bothered railway enterprises in the daily operation. Statistics show that the failures also increase with the PM action frequencies, and thus there is a critical requirement to investigate the reliability and PM policy of combined governors.

At present, various approaches have been devoted to improve the maintenance strategy for key components of diesel locomotives, in additional to improving their function. Some of them focused on condition-based maintenance, where the maintenance duration or some aided decision are made via collecting actual technical state of key components based on monitoring [1, 4] or detecting information [1, 13]. Though potential failures of certain key components may be detected in time by this mode, it has not gained wide application to most of Chinese diesel locomotives as only a small number of components which can be checked. Only few of them determine maintenance according to some parameters of key components. Lingaitis et al, [12] proposed a method to determine the maintenance data using the state of fuel consumption of diesel locomotives, while the fuel consumption of diesel locomotive is easily influenced by some unpredictable rand factors, such as state of railway and traction weight and outside condition and so forth. Zhang et al, [22] considered the influence of environmental condition to diesel engine system of diesel locomotive and optimized the PM interval in different seasons, while the reliability does not consider the sequential PM and the maintenance limitation of the whole machine in that paper. Di [2] employed a physical model based on calculating the accumulative damage degrees of main generator according to plenty of operation records, and then determined their major maintenance period, and yet the physical model is restricted more because of the complexity of its failure mechanism. While it is difficult to monitor or detect the working state of the combined governor, and thus the PM is seen as an effective approach to improve the reliability of component in reality.

However, for the combined governor, PM activities must obey the laws of the PM Criterion of diesel locomotives. The current PM interval in diesel locomotives is decided by the operational state of the whole locomotive, and a PM action is performed when running-mileage of locomotives reached a fixed interval (23000 km ~25000 km). Combined governors always passively replaced or maintained when failures occurs, and failures usually cause large economic loss. Thus, in order to hold a high operational reliability and meet the requirements of the reality, a CR policy using the preventively maintained spare component in the PM process is discussed in this paper. Meanwhile, authentic failure data on the combined governors of DF4 diesel locomotives from one Chinese railway bureau are taken as an example, from which reliability models on four sequential operational phases are obtained. In a word, the present study as a whole falls into two main parts: one is on the modeling of lifetime distribution for four sequential operational phases using real data, the other is on the maintenance optimization of the CR policy based on obtained lifetime distribution models and the PM Criterion of diesel locomotives, and then four optimization results are obtained for the corresponding operational phases using the genetic algorithm.

The rest of the paper are organized as follows: some related works are reviewed in section 2, the process of building four reliability models of combined governors are introduced in section 3, the maintenance optimization is presented in section 4, and then a brief summary is given in the last section.

Prior to providing a detailed description on maintenance policy, some terminologies that are widely used in the forthcoming sections are introduced as follows:

$F_i(t)$ The failure function before the i_{th} PM action

$R_i(t)$ The reliability function before the i_{th} PM action

$h_i(t)$ The hazard rate function before the i_{th} PM action

$H_i(t) = \int_0^t h_i(x)dx$ Cumulative failure rate before the ith PM action

$$\delta_{i,j,k} = \begin{cases} 1 & \textit{if the component of the group } j \textit{ is selected at the kth replacement} \\ & \textit{in the ith operational phase} \\ 0 & \textit{else} \end{cases}$$

$$S_i(X) = g_i^2(X) \qquad i = 1, 2, \cdots, m$$

$R(X)$ Penalty factor

$V(X)$ Objective function

$P(X)$ Penalty function

$U(X)$ Modified objective function

T The fixed maintenance period

C_p PM cost

C_f Minimal repair cost

C_{Re} Replace cost

C_{st} Set-up cost

2. Related works

For the modeling of general lifetime distribution in maintenance fields, complete data and censored data are attached more importance than zero-failure data. These observations are divided into two categories according to their source: one is the testing data from laboratories, the other from some practical occasion. Both include complete data, censored data and zero-failure data. Models on lifetime distribution with complete data or censored data are more mature than those uniform distribution on the domain (0,1) and (1,c). Let c=4 in this case.

c) Set the likelihood function of p_i as $L(0 \mid p_i) = (1-p_i)^{s_i}$. According to the Bayesian theorem, the posterior density function of p_i is:

$$h(p_i \mid s_i) = \frac{\pi(p_i \mid a,b)L(0 \mid p_i)}{\int_0^1 \pi(p_i \mid a,b)L(0 \mid p_i)dp_i}$$

d) Using quadratic loss function gives the Bayesian estimation of p_i as:

$$\hat{p}_{iB} = \int_0^1 p_i h(p_i \mid s_i)dp_i$$

e) Calculate the E-Bayes estimation of p_i is

$$\hat{p}_{iEB} = \iint_D \hat{p}_{iB}(a,b)\pi(a,b)dadb = \frac{1}{3}\int_1^4\int_0^1 \frac{a}{s_i+a+b}dadb \tag{1}$$

Then, Each t_i versus reliability $R(t_i)$ is converted into a new axis by the Weibull Conversion and plot the Weibull Probability Plot (WPP), as shown in Fig. 1, from which can find that the distribution tendency is a concave curve with monotone increase, and which is befitting

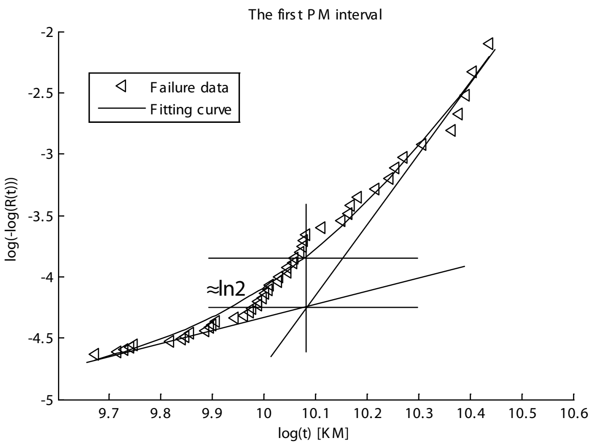


Fig. 1. WPP Plot for the 1st operational phase

for the two-fold Weibull distribution of a competing risk model. The initial parameters are estimated by graphic method, reader can refer literature [8]. The further parameters estimation is made by the least square method using the Matlab 2010b. The reliability distribution function for the first operational phase is determined as follows:

$$R_1(t) = \exp(-(t/49862)^{6.19} - (t/237810)^{0.65}) \tag{2}$$

2) Reliability model for the second to forth operational phase

The statistics of other three PM intervals are made up of censored data and failure data, their initial reliability can attain from the Median Rank Estimates as formula (3).

Median Rank Estimates of initial reliability:

$$R(t_i) = 1 - \frac{i-0.3}{r_i+0.4}. \tag{3}$$

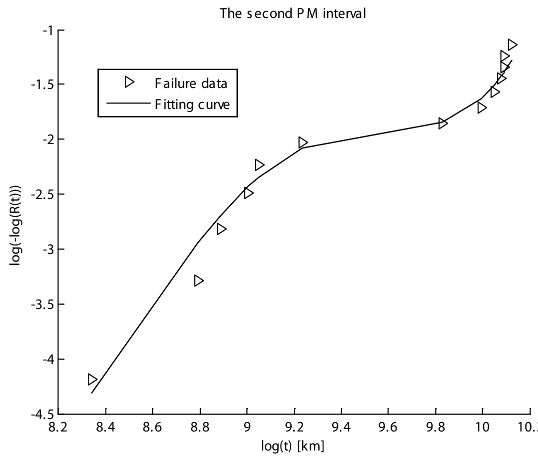
Where $r_i = r_{i-1} + \frac{n+1-r_{i-1}}{n+2-j}$

Where j is the sequence number of failure data among the whole data, and i is the sequence number of the failure data. Then every t_i is converted into versus initial reliability $R(t_i)$ utilizing the Weibull conversion and plot the WPP. As separately shown in Fig. 2 (a), (b) and (c), from which it can find that the distribution tendency are befitting with the two-fold Mixed-Weibull model.

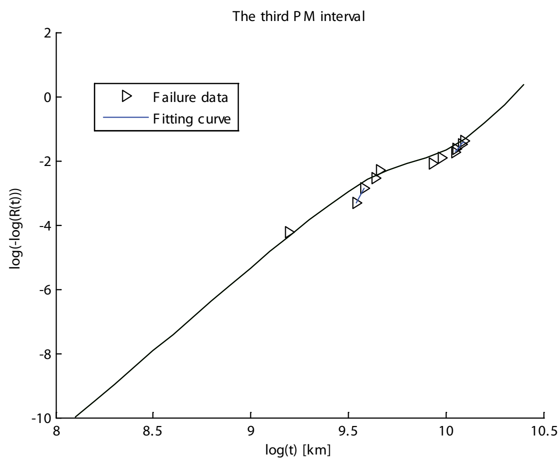
Finally, the initial parameters are estimated by graphic method, their distribution model can be determined initially from the distribution trend of the failure data, which are satisfied with the two-fold mixed-Weibull distribution. The method of their initial parameters estimation can refer literature[9], and the further parameters estimation is made by the least square method using the Matlab 2010b. It is not difficulty to find that the failure data and fitting curve have a good fit from the WPP plot.

The following are lifetime distribution functions for the second to the forth operational phase, and their WPP Plots are respectively shown as Fig.2 (a), (b) and (c).

$$R_2(t) = 0.867\exp(-(t/32919)^{7.23}) + 0.133\exp(-(t/8151)^{3.37}) \tag{4}$$

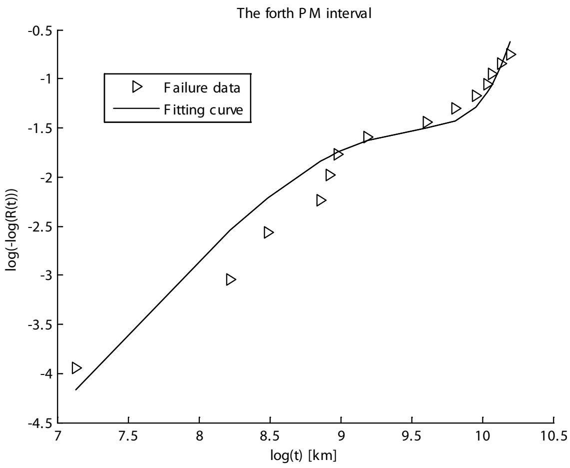


(a)



(b)

Fig. 2a, 2b. WPP Plots for the 2nd to the 4rd operational phase



(c)

Fig. 2c. WPP Plots for the 2nd to the 3rd operational phase

$$R_3(t) = 0.9\exp(-(t/31629)^{7.1}) + 0.1\exp(-(t/14732)^{5.2}) \quad (5)$$

$$R_4(t) = 0.8\exp(-(t/31488)^{7.2}) + 0.2\exp(-(t/5821)^{1.6}) \quad (6)$$

3.4. Model Analysis

It can be found that there exists large difference between the first maintenance interval and the others on reliability and hazard rate distribution from four reliability/hazard rate distribution plots. With the PM times increase, the reliability distribution is decreased rapidly when the running mileage less than 10000 km and greater than 20000 km, and it keeps relatively steady at 10000 km ~20000 km, which is benefit for users to make a maintenance decision. Meanwhile, it is a pity that the reliability is always greater than 0.9 at the first operational phase, while the combined governor must be preventively maintained according to the *PM criterion* of diesel locomotives, and thus the maintenance optimization will be discussed in section 3.

It is perfect for the first operational phase that the hazard rate is less than 1×10^{-5} when the PM action performed. For the other three hazard rate, there is an undulation near 5000 km at every hazard rate distribution, and it increase rapidly when the running mileage over 30000 km, which can be found from Fig. 3. Therefore, users should give more attention near 5000 km, and control the running mileage less than 30000 km.

4. Maintenance Optimization

Due to the importance of the combined governor in diesel locomotives, it must be preventively maintained or replaced at every PM activity of diesel locomotives in practice. Consequently, unreasonable maintenance actions maybe cause maintenance cost and down frequency increase. One effective approach is to replace the unstable operating component with a spare one in the midway, where the spare component is preventively maintained and without service after the PM action. Thus, how long the replacement should be performed and how to combine the operating component with a spare one must be determined. Therefore, in this section, a CR policy is discussed in detail.

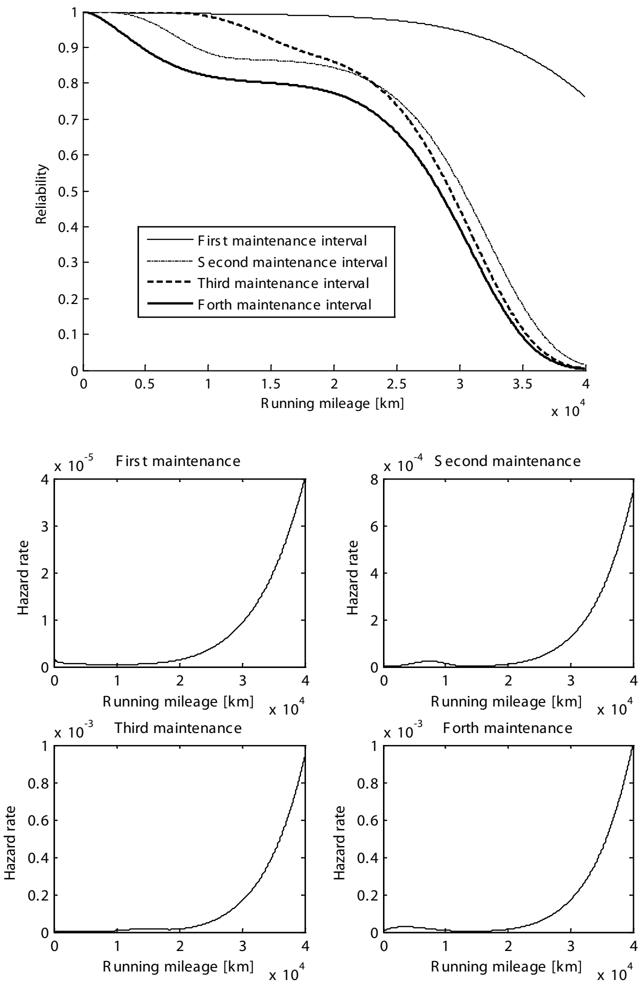


Fig. 3. Reliability/Hazard rate distribution

4.1. Analysis of Existing PM Approach of Combined Governors

As mentioned above, the existing maintenance criterion of diesel locomotives in China includes periodic PM and overhauling. The upper limit of the periodic maintenance interval is not a fixed value, but it is usually a domain such as 23000 km~25000 km. During the overhauling period, three periodic PM actions are performed, and almost each main component must be maintained at a PM process, which may cause some components being preventively replaced in the midway because of the under/over-maintenance, the combined governor is a case in point. Once its operational condition is unsatisfied the operational requirement or malfunctions take place frequently, a replacement or a minimal repair must be done. The spare combined governor which is preventively maintained (<4) is random selected if need to be replaced, and it would be scrap or instead by a new one after a fixed maintenance activities. Take into consideration its cost, it is impossible to replace an unstable one using a new one for enterprises, while preventively maintained component is always performed to replace the unstable one. Meanwhile, the combined governor must be preventively maintained if an installation and take-down take place.

4.2. The CR Policy

According to the current maintenance condition and obtained reliability models, a CR policy is proposed for the sequential PM of combined governors under a given maintenance interval of diesel locomotives.

The CR policy is based on the maintenance criterion of diesel locomotives that three PM actions are performed in a periodic of an overhauling which is divided into four operational phases, at the end of the last operational phase, a overhauling or renewal is done. During each operational process, malfunctions are removed by minimal repair, and the CR policy is performed under the following assumptions:

1) During an operation period, the finite CR action is performed using spare components which are preventively maintained at least once and without service after PM. Spare components are sufficient and in cold standby state, and can be selected randomly if needed. New components are only used at the first operational phase.
2) The minimal repair time and set-up time can be ignored, while the repair cost and the set-up cost exist and are regarded as a fixed constant.
3) A component which undergoes the same number of PM activities has the same service time in one operational phase.

The CR process is illustrated as follows:

In the *i*th operational phase, a CR action occurs at the time $t_{i,j}$ (i,j =1,2,…,n), in which a spare component which is in cold standby state after PM is selected to replace the failed one. The CR process in one operational phase is indicated as Fig. 4, where $x_{i,j}$ denotes the operational interval of the component which is underwent *j*-1 PM actions (belongs to the group *j*) and served in the i^{th} operational phase, and the number *j* (*j=2,3,4*) denotes the component which belongs to group *j*, and "no" denotes no replacement. Components used in a CR process in one operational phase is represented as (a_1 a_2 … a_n), where the value of a_j (*j*=1,2,…,n) denotes the accumulative using number of components which belong to the *j*th group used in the ith operational phase. When *j*=1, the selected component is new, such as (1 0 0 2) represents that three components are applied in this operational phase, one is new and the other two belong to the group 4 and are employed thrice PM action.

According the maintenance theory, the replacement times may be infinite, while for real maintenance processes, the system cannot be performed by repair actions all along for minimizing the long-term total cost. That is to say, repair actions cannot be chosen without any restriction, because there will never be an infinite number of repairs in finite time, and thus the replacement times less than 4 in this paper.

Based on the maintenance cost, the mean cost rate in the ith operational phase can be show as the follows:

$$C_i(x_{i,i},x_{i,j})=\frac{C_f(H_i(x_{i,i})+\sum_{j=2}^{m_0}\sum_{k=0}^{K}\delta_{i,j,k}H_j(x_{i,j}))+(\sum_{j=2}^{m_0}\sum_{k=0}^{K}\delta_{i,j,k}+1)(C_{st}+(C_p\,|\,C_r))}{\sum_{\substack{j=2\\ j\neq i}}^{m_0}\sum_{k=0}^{K}\delta_{i,j,k}x_{i,j}+x_{i,i}} \tag{7}$$

$$\text{St.}\quad \begin{aligned}&\sum_{j=2}^{m_0}\sum_{k=0}^{K}\delta_{i,j,k}+1\le N_0\\ &T_{lo}\le\sum_{j=2}^{m_0}\sum_{k=0}^{K}\delta_{i,j,k}x_{i,j}+x_{i,i}\le T_{up}\end{aligned} \tag{8}$$

$$\text{Where }\ \delta_{i,j,k}=\begin{cases}1 & \textit{if the component of the group } j \textit{ is selected at the kth replacement}\\ & \textit{in the ith operational phase}\\ 0 & \textit{else}\end{cases}$$

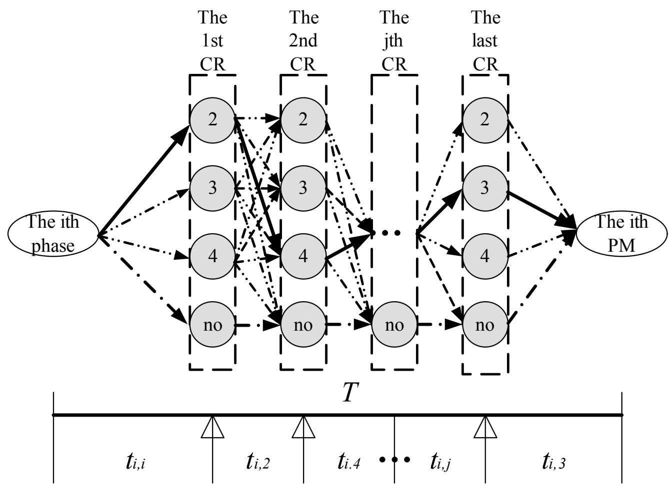


Fig. 4. CR process

$H_i(x_{ii})$ is accumulative failure times of the component which is suffered *i*-1 PM actions and employed in the *i*th operational phase, $x_{i,i}$ is its operational time, and $x_{i,j}$ is the same as above. $\delta_{i,j,k}$ is discussed in the section 3.4. ($C_p|C_r$) means C_r or C_p, the cost is C_r if the operational component belong to the last group, and else it is C_p. The T_{lo} and T_{up} are lower limit and upper limit of the maintenance interval, and N_0 is the allowable accumulative replacement times in an operational phases. In this paper the replacement times is less than 4, and the operational phase number is 4, T_{lo}=23000 km and T_{up}=25000 km.

4.3. Solution of the maintenance model

The proposed maintenance model is a constraint optimization problem which is the most important and ubiquitous type of engineering optimization problems. Evolutionary algorithms (EA) have been applied extensively for tackling these problems with various degrees of success. The penalty function approach is a relatively simple approach and is remarkably easy to implement and, as a result, has been popularly used with an EA [6]. In this paper, the genetic algorithm (GA) is applied with the penalty function approach. The constrained objective function Eq. 8 can be converted into an unconstrained one by adding the penalty function. Then, the equivalent unconstrained optimization problem can be stated as:

$$\text{Min: } U(X)=V(X)+P(X) \tag{9}$$

Where *V*(*X*) is the objective function, *P*(*X*) is the penalty function and *U*(*X*) is the modified objective function which is constrained by the function $g_i(X)$. The penalty function *P*(*X*) is defined such as:

$$P(X)=\begin{cases}0 & X\in R^n\\ \sum_i S_i(X)R_i(X) & X\notin R^n\end{cases} \tag{10}$$

Where $S_i(X)$ is the real-valued continuous function defined by

$$S_i(X)=g_i^2(X)\qquad i=1,2,\cdots,m \tag{11}$$

and $R_i(X)$ is penalty factor, which is stepwise changes with the violation for each of constraint and is indicated by:

$$R_1(X)=\begin{cases}0 & 23000\le X\le 25000\\ 200 & 1000<|X-24000|\le 1500\\ 280 & 1500<|X-24000|\le 2000\\ 350 & 2000<|X-24000|\le 2500\\ 500 & else\end{cases} \quad (12)$$

Then, the determined function $U(X)$ has a high-dimension, and thus the genetic algorithm is suitable for this case. Let C_f=8000, C_p=1000, C_{re}=400, C_{st}=300, and the variable t belongs to [500 25000] on the basis of the practice. The GA's parameters are chosen as follows:

Chromosome length=20, population size=400, crossover probability=0.85, mutation probability=0.02. Using the Matlab 2010b the optimization is obtained shown as Table 2.

4.4. Optimization Result

Optimization results are shown in Table 2, it can be seen that optimized replacement policy exists in every operational phase, and the optimal results are marked with grey. The determined policy for the first phase is (1 0 0 0), the PM interval is 25000 km, and the minimal cost rate is 0.1305. For the second phase, the results respectively are: (0 1 0 0), 23000 km and 0.1322, the third phase are: (0 0 2 0), 25000 km and 0.1274, and the forth phase are: (0 0 1 1), 25000 km and 0.14361. It can be seen from the first two phases that the existing maintenance method is reasonable and no CR is needed, but the combination (1 0 1 0) in the first phase and (0 1 1 0) in the second phase are regarded as optional policy if spare components are limited. The reasonable combined CR can decrease the cost rate at the last two phases, and thus the existing maintenance approach should be modified. Moreover, the others optimization results are useful, which can be viewed as a dynamic conference for users when the replacement is limited by spare components.

4.5. Discussion

1) $\delta_{i,j,k}$

$\delta_{i,j,k}$ is the indicator function, it is 1 if a component of the group *j* is selected at the *k*th replacement in the *i*th operational phase, and else it is 0. In order to convenient discussion, let $\sum_{k=0}^{n}\delta_{i,j,k}=n_{i,j}$, the $n_{i,j}$ is accumulative total number of a component with *j*-1 PM actions used in the ith operational phase, which is a permutation and combination value decided by the *i* and the allowable CR times *n*. Let *m* denotes the maintenance phase and *n* denotes the allowable CR times. The permutation and combination value N_{one} in one operational phase and the total number N_{total} indicates as follows:

$$N_{\text{one}}=(\text{m}-1)^{n-1}+1 \quad (11)$$

$$N_{total}=3N_{one}-1 \quad (12)$$

It can be seen that with the increase of *m* and *n*, N_{one} and N_{total} are with exponential growth. In reality of railway enterprises, the value *m* usually is 4~6, and *n* is 2~4. Thereupon, the enumerate approach can be used. For instance, in this paper, *m*=4, and *n*=3, N_{one} is 10 for one maintenance phase, and the total number is 29 which can be seen in Table 2.

2) Comparative on the CR policy and (T,δ) policy

In the section 1, the policy that Tango proposed and Sheu proposed are introduced briefly. Tango held if items fail in [$(k-1)T$, $kT-v$), they are replaced by new items, and if in [$kT-v$, kT), they are replaced by used items. Sheu considered that an operating system can be preventively replaced by new ones at times iT ($i=1,2,\ldots$) independently of its failure history, and if system fails in ($(i-1)T$, $iT-\delta$) it is either replaced by a new one or minimally repaired, and if in [$iT-\delta$, iT) it is either replaced by a used one or minimally repaired. The parameters v and δ can be seen as a threshold for these policies, in which one maintenance approach is performed if the lifetime less than the threshold otherwise another maintenance method is made. We also assume that failures are removed by minimal repair if the components fail in (0, δ] ($0<\delta<T$), otherwise, the component is preventively replaced by a component that is employed PM actions at least once in (δ, T), and which is preventively maintained or replaced at T. This policy is similar with Tango's and Sheu's and marked as (T,δ) policy. Table 3 (a) and Table 3 (b) is separately show the optimal results of δ=18000 km and δ=20000 km, where italics are policies with minimal cost rate, which proved that CR policies proposed this paper are more reasonable than (T,δ) policy.

3) The optimization results

According to the PM *criterion of diesel locomotives* that the locomotives' PM action is performed in the interval of 23000 km ~25000 km instead of a fixed value, which is reasonable because the PM action always is effected by some dynamic factors in the reality. Optimization results for CR policy are shown in Table 2, which illustrated that the existing maintenance approach is the same as proposed CR policy in the first two operational phases and no CR is needed, while it should be replaced by the proposed policy in the last two operational phases. According to the common consideration, the (T,δ) policy is easily accepted by ordinary costumers, while the compare in Table 4 show that the proposed CR policy is more reasonable than others. In addition, the rest optimization results of proposed CR can be regard as a conference to make a replacement decision if spare components are finite.

5. Conclusions

It is widespread that subcomponents must follow the PM schedule of the whole machine, just as combined governors of diesel locomotives which must be preventively maintained together with the diesel locomotive, this maintenance policy may cause over/under-maintenance and leads to passive replacement/maintenance and large economic loss in practice. How to determine an optimal policy considered the reality of railway enterprises, which includes compare components and preventive maintenance criterion of diesel locomotives, is much helpful to railway enterprises. Therefore, four reliability models for combined governors are obtained via making the most use of the real data from one Chinese Railway Bureau. Furthermore, a novel CR policy is also presented under sequential PM of diesel locomotives, in which replaced components apply used items instead of new items for the consideration of cost in practice, and the upper limit of the PM interval is not a fixed value, but it is in 23000 km ~25000 km. The results demonstrate that:

1) The reliability of combined governors in the first phase is very reliable, indicating that no PM action is needed which is the same as the optimized results. Then, the reliability decreases rapidly when the lifetime over 20000 km and there is an undulation of hazard rate in 5000 km ~10000 km after a PM performed at the rest operational phases, suggesting the users should keep in mind. Obtained reliability models of combined governors can be used to in grouped maintenance and performance improvement in diesel locomotives.
2) The optimal results show that the proposed CR policy is the best policy among the existing policy and the (T,δ) policy. It is the same as the existing maintenance approach for the first two operational intervals, while the current maintenance policy

Table 2. Optimization results on every operational phase

Phase	Group 1	Group 2	Group 3	Group 4	Cost rate	CR policy				Interval
the 1st	24998.51	0	0	0	0.130467	1	0	0	***0***	***24999***
	24499.18	0	0	500.8065	0.182226	1	0	0	1	25000
	23999.8	0	0	500.0895	0.234032	1	0	0	2	25000
	14520.25	0	10479.68	0	0.161302	1	0	1	0	25000
	14238.72	0	10261.16	500.1194	0.213373	1	0	1	1	25000
	500	0	12249.4	0	0.183232	1	0	2	0	24999
	21384.31	3615.678	0	0	0.175229	1	1	0	0	25000
	20858.58	3641.236	0	500.1194	0.227228	1	1	0	1	25000
	10799.76	3605.548	10592.89	0	0.207013	1	1	1	0	24998
	18420.85	3289.464	0	0	0.221275	1	2	0	0	25000
the 2nd	***0***	***23001.13***	***0***	***0***	***0.132197***	***0***	***1***	***0***	***0***	***23001***
	0	22921.64	0	1628.198	0.184266	0	1	0	1	24550
	0	22478.55	0	1260.724	0.232655	0	1	0	2	25000
	0	16594.63	8405.362	0	0.153641	0	1	1	0	25000
	0	4816.213	19571.46	612.328	0.208291	0	1	1	1	25000
	0	3668.288	10665.85	0	0.170046	0	1	2	0	25000
	0	12499.81	0	0	0.194488	0	2	0	0	25000
	0	3526.155	0	17947.61	0.237732	0	2	0	1	25000
	0	5269.242	14461.43	0	0.194598	0	2	1	0	25000
	0	8332.923	0	0	0.24418	0	3	0	0	24999
the 3rd	0	0	23001.13	0	0.1294	0	0	1	0	23001
	0	0	21667	1536.598	0.179661	0	0	1	1	23204
	0	0	21748.63	1625.673	0.227887	0	0	1	2	25000
	0	16460.45	8539.382	0	0.153657	0	1	1	0	25000
	0	16566.91	7932.964	500.1194	0.206427	0	1	1	1	25000
	0	5229.446	14541.07	0	0.194598	0	2	1	0	25000
	0	3712.789	10643.59	0	0.170046	0	1	2	0	25000
	0	***0***	***12499.81***	***0***	***0.127395***	***0***	***0***	***2***	***0***	***25000***
	0	0	12249.4	501.1953	0.178659	0	0	2	1	25000
	0	0	8332.923	0	0.160921	0	0	3	0	24999
the 4th	0	0	0	23001.13	0.144281	0	0	0	1	23001
	0	0	0	12499.81	0.218299	0	0	0	2	25000
	0	0	0	8332.923	0.306472	0	0	0	3	24999
	0	***0***	***21566.4***	***3433.71***	***0.14361***	***0***	***0***	***1***	***1***	***25000***
	0	4836.668	19567.95	595.3517	0.18429	0	1	1	1	25000
	0	0	12247.21	505.5824	0.154659	0	0	2	1	25000
	0	3524.05	0	17951.89	0.213732	0	2	0	1	25000
	0	22392.84	0	1191.86	0.159329	0	1	0	1	23585
	0	0	21752.53	1623.733	0.203887	0	0	1	2	25000
	0	22497.34	0	1251.278	0.208655	0	1	0	2	25000

Table 3 (a). δ=18000

Phase	Group 1	Group 2	Group 3	Group 4	Cost rate	CR policy				Interval
the 1st	***24999.87***	***0***	***0***	***0***	***0.130464***	***1***	***0***	***0***	***0***	***25000***
	21481.41	3518.585	0	0	0.175246	1	1	0	0	25000
	18000	0	6999.991	0	0.165028	1	0	1	0	25000
	24499.92	0	0	500.0597	0.182225	1	0	0	1	25000
the 2nd	***0***	***23000.03***	***0***	***0***	***0.132195***	***0***	***1***	***0***	***0***	***23000***
	0	18000	6999.991	0	0.154408	0	1	1	0	25000
	0	22921.88	0	1628.198	0.184266	0	1	0	1	24550
the 3rd	0	0	***23000.03***	***0***	***0.129394***	***0***	***0***	***1***	***0***	***23000***
	0	0	21666.44	1536.598	0.179661	0	0	1	1	23203
	0	5014.835	19985.08	0	0.157342	0	1	1	0	25000
the 4th	0	0	0	23000.03	0.144276	0	0	0	1	***23000***
	0	4898.265	0	20101.72	0.171069	0	1	0	1	25000
	0	0	6999.155	18000.81	0.154599	0	0	1	1	25000

Table 3 (b). δ=20000

Phase	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Cost rate**	**CR policy**				**Interval**
the 1st	***24999.91***	***0***	***0***	***0***	***0.130464***	***1***	***0***	***0***	***0***	***25000***
	21379.12	3620.861	0	0	0.175228	1	1	0	0	25000
	20000.07	0	4999.889	0	0.16925	1	0	1	0	25000
	24499.86	0	0	500.1194	0.182225	1	0	0	1	25000
the 2nd	***0***	***23000.03***	***0***	***0***	***0.132195***	***0***	***1***	***0***	***0***	***23000***
	0	20000.07	4999.889	0	0.158505	0	1	1	0	25000
	0	22921.85	0	1628.198	0.184266	0	1	0	1	24550
the 3rd	***0***	***0***	***23000.03***	***0***	***0.129394***	***0***	***0***	***1***	***0***	***23000***
	0	0	21666.47	1536.598	0.179661	0	0	1	1	23203
	0	4998.397	20001.57	0	0.157343	0	1	1	0	25000
the 4th	***0***	***0***	***0***	***23000.03***	***0.144276***	***0***	***0***	***0***	***1***	***23000***
	0	4886.292	0	20113.71	0.17107	0	1	0	1	25000
	0	0	4999.889	20000.07	0.163649	0	0	1	1	25000

Table 4. Results

Phase	**the 1st CR**	**the 2nd CR**	**the 3rd CR**	**the 4th CR**	**Cost rate**	**CR policy**				**Interval**	**policy**
the 1st	24998.51	0	0	0	0.130467	1	0	0	0	24999	same
the 2nd	0	23001.13	0	0	0.132197	0	1	0	0	23001	same
the 3rd	0	0	23001.13	0	0.12941	0	0	1	0	23001	existing
	0	*0*	*12499.81*	*0*	*0.127395*	*0*	*0*	*2*	*0*	*25000*	*CR*
	0	0	23000.03	0	0.129394	0	0	1	0	23000	*δ=18000*
	0	0	23000.03	0	0.129394	0	0	1	0	23000	*δ=20000*
the 4th	0	0	0	23001.13	0.144281	0	0	0	1	23001	existing
	0	*0*	*21566.35*	*3433.712*	*0.14361*	*0*	*0*	*1*	*1*	*25000*	*CR*
	0	0	0	23000.03	0.144276	0	0	0	1	23000	*δ=18000*
	0	0	0	23000.03	0.144276	0	0	0	1	23000	*δ=20000*

need to be modified with (0 0 2 0) and (0 0 1 1) in the last two operational phases. Meanwhile, the others combinatorial replacement results applying maintained items can help users to make a decision when spare components are limited.

In the future working, we will further investigate the application of proposed models into improvement of the maintenance effect of diesel locomotives. Further, since spare components are critical factor in maintenance, the effects of spare components number on the maintenance of locomotives is also very interesting and promising work.

Acknowledgement

The research work is supported by the Natural Science Foundation of China (51275090, and 71201025), the Hexa-Type Elites Peak Program of Jiangsu Province (2008144), and the Innovation Fund for Graduate Students of Jiangsu Province (CXLX12_0078), China.

References

1. Albarbar A, et al. Diesel engine fuel injection monitoring using acoustic measurements and independent component analysis. Measurement. 43(10): 1376–1386.
2. DI W. Repair Period of Diesel Locomotives Based on Accumulative Damage Degree. China Railway Science. 30(6): 104-107.
3. F. M. H. and W. R. A. A Bayesian Zero-Failure(BAZE) Reliability Demonstrate Testing Procedure. Journal of Quality Technology. 11(3): 128–137.
4. Fry K. N. and BSc. Diesel locomotive a reliability improvement by system monitoring. Proc IMechE: Part F. 209(1): 1–10.
5. HAN M. E-Bayesian Estimation of the Products Reliability when Testing Reveals no Failure. Chin. Quart. J. of Math. 24(3): 407–414.
6. Homaifar A et al. Constrained Via Optimization Genetic Algorithms. SIMULATION 62(4): 242–254.
7. Jiang P et al. Reliability Estimation in a Weibull Lifetime Distribution with Zero-failure Field Data. Qual. Reliab. Engng. Int. 26(7): 691–701.
8. Jiang R, Murthy DNP. Study of n-Fold Weibull Competing Risk Model. Mathematical and Computer Modelling. 38(11): 1259–1273.
9. Jiang R, Murthy DNP. Reliability modeling involving two Weibull distributions. Reliability Engineering and System Safety. 47(3): 187–198.
10. Lai KK et al. Practices of preventive maintenance and replacement for engines: A case study. European Journal of Operational Research. 124(2000): 294–306.
11. Lhorente B et al. A model for optimal armature maintenance in electric haul truck wheel motors: a case study. Reliability Engineering and System Safety. 84(2004): 209–218.
12. Lingaitis LP et al. Prediction Methodology of Durability of Locomotives Diesel Engines. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (2): 154–159.
13. Macian V et al. Results and benefits of an oil analysis programme for railway locomotive diesel engines. Insight. Jun; 45(6): 1–5.
14. Martz HFJ, Waller RA. Bayesian Zero-Failure (BAZE) reliability demonstration testing procedure. J. Qual. Technol. 11(3)(128–137.
15. Miller KW et al. Estimating the Probability of Failure When Testing Reveals No Failures. IEEE TRANSACTIONS ON SOFTWARE ENGINEERING. 18(1): 33–43.
16. RE. B. and H. LC. Optimal preventive maintenance policies. Operations Research. 8(1): 90–100.
17. Sheu S-H, Griffith WS. Extended block replacement policy with shock models and used items. European Journal of Operational Research. 140(2002): 50–60.
18. Tango T. Extended Block Replacement Policy with Used Items J. Appl. Prob. 15(3): 560–572.
19. Wang H. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research. 139(2002): 469–489.
20. Weber P. and L. Jouffe. Complex system reliability modelling with Dynamic Object Oriented Bayesian Networks (DOOBN). Reliability Engineering and System Safety. 91(2006): 149–162.
21. Zhang T, MinXie. On the upper truncated Weibull distribution and its reliability implications. Reliability Engineering and System Safety. 96(2011): 194–200.
22. Zhang Z et al. Reliability Modeling and Maintenance Optimization of the Diesel System in Locomotives. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 302–311.
23. Zhao X et al. Optimal imperfect preventive maintenance policies for a used system. International Journal of Systems Science. 43(9): 1632–1641.

Wenke GAO, Ph.D. Candidate
Prof. Zhisheng ZHANG
Dr. Yifan ZHOU
School of Mechanical Engineering
Southeast University
Nanjing, China, 211189
E-mail: oldbc@seu.edu.cn
Eng. Rujiao WANG
Tianjin Locomotive Maintenance Corporation
Tianjin, China, 301700
Eng. Lei JIANG
Freight Department of Lanzhou Railway Bureau
Lanzhou, China, 730030

Article citation info:
FUCHS P, ZAJICEK J. Safety integrity level (SIL) versus full quantitative risk value. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 99–105.

Pavel FUCHS
Jaroslav ZAJICEK

SAFETY INTEGRITY LEVEL (SIL) VERSUS FULL QUANTITATIVE RISK VALUE

NIENARUSZALNOŚĆ BEZPIECZEŃSTWA A WARTOŚĆ RYZYKA

Safety management of technical equipment is not possible without risk assessment. Therefore, many standards are available for risk assessment, e.g. ISO 13824:2009 General principles on risk assessment of systems involving structures or ISO/IEC 31010:2009 Risk management – Risk Assessment Techniques. In different industrial sectors risk assessment is fundamental step to determine required safety integrity level (SIL), eventually performance level (PL), which guarantees risk linked to some equipment on acceptable level. Standards applied for risk management based on SIL in different industrial sectors differ in methods used for risk evaluation and SIL determination. IEC 61508-5 accepts the use of qualitative, semi-quantitative or quantitative approach for risk evaluation and SIL determination. The standard uses hazardous event severity matrix as an example of qualitative approach for SIL determination, the standard furthermore uses layer of protection analysis (LOPA) as an example of semi-quantitative approach. The standard also uses Risk graph method as an example of both qualitative and semi-quantitative approach. IEC 62061 only presents one semi-quantitative approach for risk evaluation and SIL determination based on combination of probability and severity of consequences. This approach is different from the approach presented in IEC 61508-5. Similarly ISO 13849-1 recommends the use of qualitative method combining probability and severity of consequences for risk evaluation and PL determination, however, distinctly from IEC 61508-5 as well as IEC 62061. All these standards evaluate risk in the first step and in the second step they set safety systems reliability requirements, which should lower risk onto an acceptable level. The elemental question is, how exactly these standards evaluate risk in their methods. Another question is what acceptable level of risk is implicitly hidden in their requirements for choice of SIL and PL. This paper addresses these questions.

***Keywords**: safety, SIL and PL determination, risk evaluation, tolerable level of risk, semiquantitative analysis.*

Zarządzanie bezpieczeństwem urządzeń technicznych nie jest możliwe bez oceny ryzyka. Dlatego też istnieje wiele norm związanych z oceną ryzyka, np. ISO 13824:2009 Ogólne zasady dotyczące oceny ryzyka w systemach obejmujących konstrukcje lub ISO/IEC 31010:2009 Zarządzanie ryzykiem - Techniki oceny ryzyka. W różnych gałęziach przemysłu ocena ryzyka jest podstawowym krokiem na drodze do określenia wymaganego poziomu nienaruszalności bezpieczeństwa (SIL), oraz ewentualnie poziomu wydajności (PL), który gwarantuje, że ryzyko w odniesieniu do niektórych urządzeń pozostanie na akceptowalnym poziomie. Normy stosowane w zakresie zarządzania ryzykiem w oparciu o SIL w różnych gałęziach przemysłu różnią się jeśli chodzi o metody stosowane do oceny ryzyka i określenia SIL. IEC 61508-5 akceptuje zastosowanie jakościowego, pół-ilościowego lub ilościowego podejścia do oceny ryzyka oraz określenia SIL. Norma ta wykorzystuje macierz ciężkości zdarzeń niebezpiecznych jako przykład podejścia jakościowego do określenia SIL; ponadto, norma wykorzystuje analizę warstw zabezpieczeń (LOPA) jako przykład podejścia półilościowego. Norma wykorzystuje również metodę wykresu ryzyka jako przykład podejścia zarówno jakościowego jak i półilościowego. IEC 62061 prezentuje jedno pół-ilościowe podejście do oceny ryzyka i określenia SIL łącząc prawdopodobieństwo i ciężkość następstw. To podejście różni się od metody stosowanej w IEC 61508-5. Podobnie ISO 13849-1 zaleca stosowanie metody jakościowej łączącej prawdopodobieństwo i ciężkość następstw dla oceny ryzyka i określenia PL, jednak w sposób odmienny od IEC 61508-5 oraz IEC 62061. Wszystkie powyższe normy dokonują oceny ryzyka w pierwszym etapie zaś w drugim etapie ustalają one wymagania odnośnie niezawodności systemów bezpieczeństwa, które powinny obniżyć ryzyko do akceptowalnego poziomu. Podstawowym pytaniem jest jak dokładnie powyższe normy dokonują oceny ryzyka przy użyciu swoich metod. Inną kwestią jest to, jaki dopuszczalny poziom ryzyka jest domyślnie ukryty w ramach ich wymagań dotyczących wyboru SIL i PL. Niniejszy artykuł odnosi się do powyższych zagadnień.

***Słowa kluczoewe**: bezpieczeństwo, określenie SIL i PL, ocena ryzyka, dopuszczalny poziom ryzyka.*

1. Introduction

For effective risk management it is necessary to be able to assess the risk accordingly. Risk assessment in technical practice becomes one of the fundamental elements to prove that a piece of equipment is sufficiently safe. This can be furthermore seen in standards in various industrial sectors. These standards require risk assessment for equipment or a device and prove the risk is acceptable. The standards then offer different approach for risk evaluation; qualitative, semi-quantitative and quantitative. These standards do not generally provide specific instructions on how to proceed while assessing risk for individual cases. The standards only provide generic recommendations with respect to variety of unsafe events and their consequences. When standards these generic cases specify more in depth by using examples, they do so in the form of appendices listed as informative, hence non-obligatory.

Fundamental standards of functional safety are IEC 61508-5 [1] and IEC 61511-x [2]. Principles from these two have been adopted in various industrial sectors related to functional safety, i.e. IEC 62061 [3], ISO 13849-1 [4], IEC 61513 [5], EN 50129 [6] and more. These listed standards are the result of historical development of understanding the function of safety systems for reducing the risk resulting from operating technical equipment. When the requirements are laid down in the standards need to be applied to a specific technical solution it is necessary to appropriately understand the essence of risk and its evaluation. Examples of risk evaluation given in standards cannot be then carelessly applied. This could lead to underestimating or overestimating of the risk level and as a result safety management would be ineffective.

This article presents the analysis of simplified approaches to determining safety integrity from three international standards IEC 61508-5, IEC 62061 and ISO 13849-1. The purpose of this article is to point out some common solutions and weaknesses of risk evaluation as well as assessing its acceptability for using these standards.

2. Risk and safety integrity according to specified standards

2.1. Risk and SIL according to IEC 61508-5

IEC 61508-5 Annex E (informative) standard provides a quantitative method to determine SIL titled Risk Graph Method (Fig. 1).

This simplified procedure is based on the following equation

$$R = (f) \text{ of a specified } (C) \tag{1}$$

and assumptions $C_1 < C_2 < C_3 < C_4$; $F_1 < F_2$; $P_1 < P_2$; $W_1 < W_2 < W_3$.

where

R is the risk with no safety-related systems in place;

f is the frequency of the hazardous event with no safety-related systems in place;

C is the consequence of the hazardous event (the consequences could be related to harm associated with health and safety or harm from environmental damage).

This produces the following four risk parameters:

- consequence of the hazardous event (C);
- frequency of, and exposure time in, the hazardous zone (F);
- possibility of failing to avoid the hazardous event (P);
- probability of the unwanted occurrence (W).

Safety integrity level (SIL) of safety-related system is specified against intolerable risk by the target failure measures presented in Table 1.

Table 1. *Safety integrity levels according to IEC 61508-5 – target failure measures for a safety function*

Safety integrity level (SIL)	Average probability of a dangerous failure on demand of the safety function [1] (PFDavg)	Average frequency of a dangerous failure of the safety function [h^{-1}] (PFH)
4	≥1E-5 to <1E-4	≥1E-9 to <1E-8
3	≥1E-4 to <1E-3	≥1E-8 to <1E-7
2	≥1E-3 to <1E-2	≥1E-7 to <1E-6
1	≥1E-2 to <1E-1	≥1E-6 to <1E-5

2.2. Risk and SIL according to IEC 61508-5

IEC 62061 Annex A (informative) provides a semi-quantitative method for determining SIL. This method is based on risk matrix (Fig. 2) [7].

This simplified procedure is based on the following equation:

$$R = (Cl) \text{ of a specified } (Se) \tag{2}$$

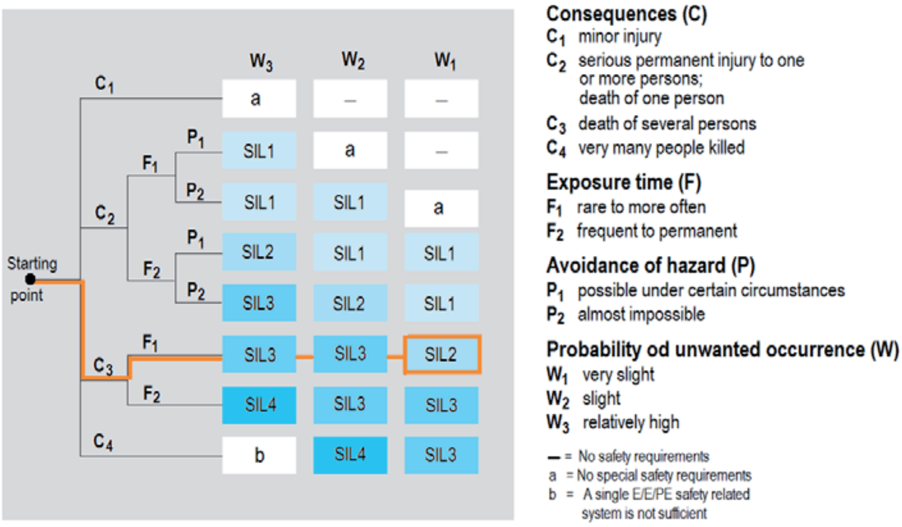


Fig. 1. The risk elements evaluation and SIL requirements determination according to IEC 61508 5

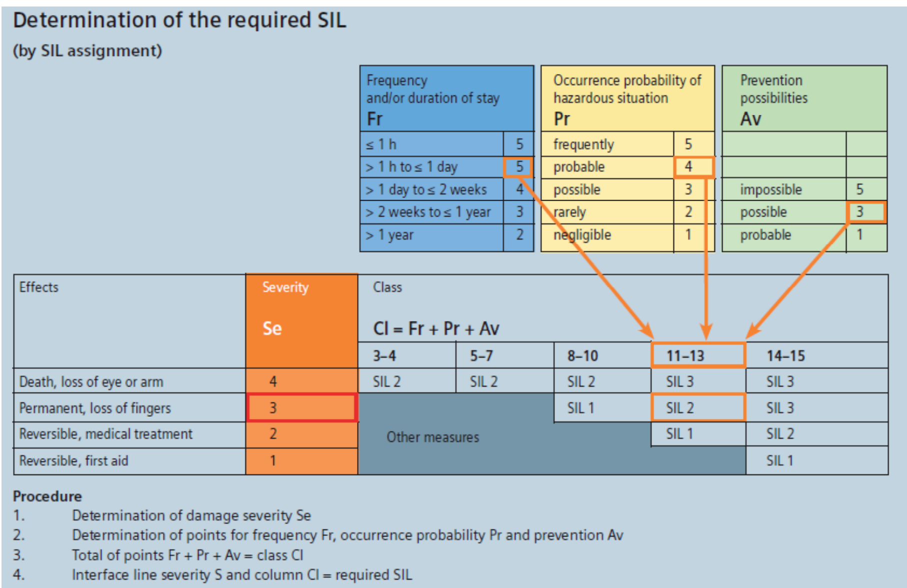


Fig. 2. The risk elements evaluation and SIL requirements determination according to IEC 62061

where

R is the risk with no safety-related systems in place;

Cl is the frequency of the hazardous event with no safety-related systems in place;

Se is the severity of consequence of the hazardous event (the consequences could be related to harm associated with health and safety).

This produces the following four risk parameters:

– consequence of the hazardous event (*Se*);
– frequency and duration of exposure to hazard (*Fr*);
– prevention possibilities (*Av*);
– occurrence probability of the hazard ous situation (*Pr*).

Safety integrity levels (SILs) of safety function according to IEC 62061 are different from SILs according to IEC 61508-5 and its target failure measures are presented in Table 2.

Table 2. Safety integrity levels according to IEC 62061 – target failure measures for a safety function

Safety integrity level (SIL)	Probability of a dangerous failures per hour [h^{-1}]
3	≥1E-8 to <1E-7
2	≥1E-7 to <1E-6
1	≥1E-6 to <1E-5

2.3. Risk and PL according to ISO 13849-1

ISO 13849-1 similarly as IEC 61508-5 in Annex A (informative) uses qualitative method based on risk graph to evaluate safety integrity, see Figure 3 [7]. With the only difference that for safety integrity of safety function the term performance level (PL) is being used.

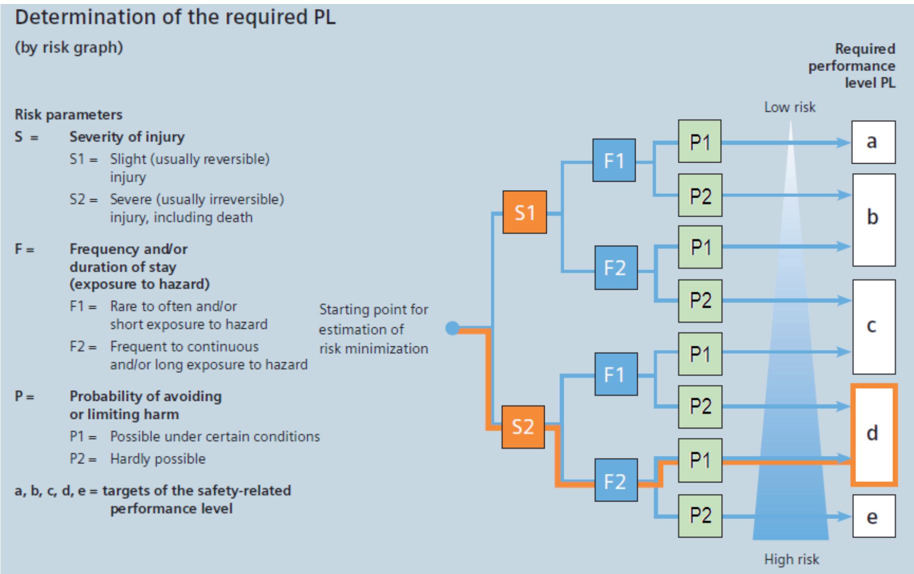


Fig. 3. The risk elements evaluation and PL requirements determination according to ISO 13849-1

This simplified procedure is based on the following equation:

$$R = (X) \text{ of a specified } (S) \quad (3)$$

where

R is the risk with no safety-related systems in place;

X is the frequency of the hazardous event with no safety-related systems in place;

S is the severity of consequence of the hazardous event (injury).

This produces the following three risk parameters:

– severity of injury (*S*);
– frequency and/or duration of exposure to hazard (*F*);
– possibility of avoiding or limiting harm (*P*).

Performance level (PL) of safety function according to ISO 13849-1 and its target failure measures are presented in Table 3.

Table 3. Safety integrity levels according to ISO 13849-1 – target failure measures for a safety function

Performance level (PL)	Average probability of a dangerous failure per hour [h^{-1}]
a	≥1E-5 to <1E-4
b	≥3E-6 to <1E-5
c	≥1E-6 to <3E-6
d	≥1E-7 to <1E-6
e	≥1E-8 to <1E-7

3. Correctness of risk evaluation and safety integrity level

3.1. Fundamental consideration of the problem addressed

Risk evaluation is linked with aleatoric uncertainty and epistemic uncertainty. Aleatoric uncertainties are given by natural randomness in the behaviour of the investigated subject. Epistemic uncertainties come from lacking knowledge of the investigated subject. The use of simplified methods then only has meaning if simplification does not radically increase epistemic uncertainties. Furthermore, only if epistemic uncertainty from knowledge of the subject's risk is not amplified by epistemic uncertainty of simplified risk evaluation.

The main purpose of this study is to find out to what extent is risk evaluation appropriate where simplified methods have been used. The result of this investigation is then better recognition of regularities valid for using simplified risk evaluation methods, hence lowering epistemic uncertainties associated with these methods.

Assessment is carried out for approaches given in IEC 61508-5, IEC 62061 and ISO 13849-1 and described in chapter 2. Their common designator is that they express the risk by the product of probabilities and consequences. Several parameters are used for it, see table 4.

It would seem that the approach according to ISO 13849-1 is different because it contains the probability (frequency) of undesirable event on an object which is a source of risk. The contradiction is only apparent. Approach according to IEC 61508-5, IEC 62061 assumes random occurrence of hazardous events in time, while ISO 13849-1 assumes the risk is permanent. The difference in approach is evident from Figure 4. In the end, both approaches evaluate risk as the product of probabilities and consequences, and determine the safety integrity level as a measure of risk reduction.

It is evident that for the quantitative risk evaluation the probability of consequence of undesirable events is given by the product of proba-

Table 4. Risk evaluation parameters according to IEC 61508-5, IEC 62061 and ISO 13849-1

Standard	Probability parameters			Consequence parameter
	Occurrence	Exposure	Avoidance	
IEC 61508-5	W	F	P	C
IEC 62061	Pr	Fr	Av	Se
ISO 13849-1	–	F	P	S

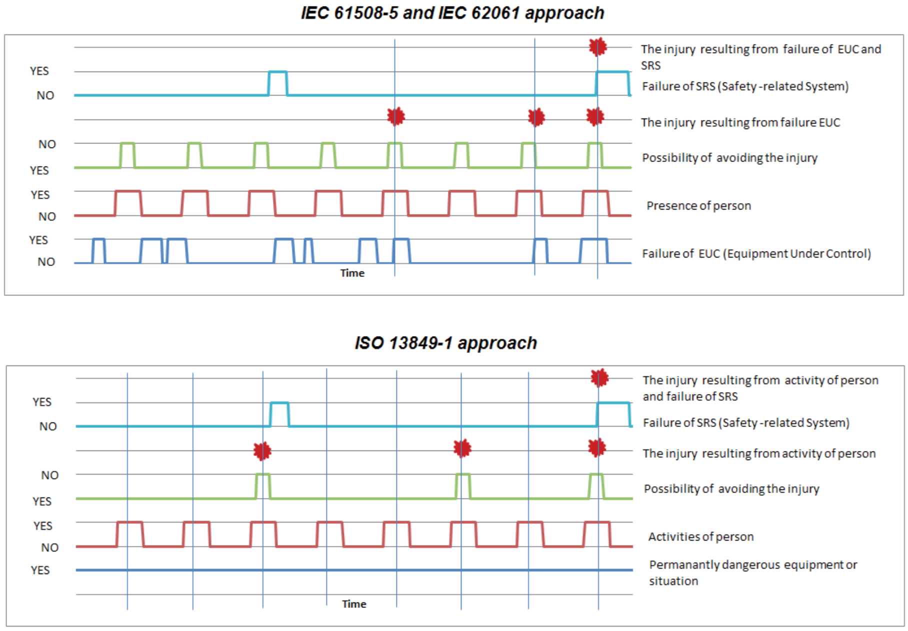


Fig. 4. The difference in dangerous situations

bility parameters and consequences. Thus, multiplying the values of all parameters takes place. If the values of these parameters are known, it is possible to evaluate the risk exactly.

The actual parameter values are not used when using the simplified approaches. Parameters are separated into zones. I instead of using the actual values, verbal (qualitative) evaluation or relative (semi-quantitative) evaluation expressed for example in points is used in these zones. Based on the given set of rules, see chapter 2, risk is then evaluated and assigned to safety integrity level. These features are common for selected standards IEC 61508-5, IEC 62061 and ISO 13849-1. Differences lie only in the evaluation parameters and rules used. Therefore, the investigation is focused on what impact on correctness lies within the choice of evaluation parameters and rules used for risk evaluation.

3.2. The assessment of the correctness of the simplified approach to IEC 61508-5

The method of determining safety integrity level (SIL) is based on a qualitative risk evaluation. Zones of parameters C, F, P and W and their ranges are assigned with verbal description. According to the description, parameter zones C, F and P can be seen as arranged in a sequence complied by a geometric scale without a specified quotient. For parameter W can be assumed that its scale is prepared according to geometric sequence with an unknown quotient.

If the simplified approach is correct, the results must match the results obtained when using fully quantitative risk evaluation. Each level of functional safety (-, a, 1, 2, 3, 4, b) is covered by successive risk intervals. With the appropriate set of parameter zones C, F, P and W, they should not overlap, see Figure 5.

Assessing the correctness of the simplified approach was based on examining whether the zone parameters C, F, P and W could be set so there was clear risk coverage through the SIL. Hence, inequality must then apply:

$$R < R_a < R_1 < R_2 < R_3 < R_4 < R_b \tag{4}$$

If the inequality is not satisfied, risk is then overlapped by two or more SIL and simplified approach cannot be considered as correct.

Using simulation in Matlab results were found for all combinations of integers in the range <2; 20> quotients scales with geometric sequence of parameters C, F, P and W. Total number of 19^4 = 130 321 possibilities were examined in accordance with the simplified approach. The observed number of overlaps of two or more SIL is shown in Table 5 and Figure 6.

If the simplified risk evaluation method was as good as the exact quantitative method, column "0" in Table 5 would have value 130 321. The distribution of non-zero values and their size suggests how sensitive is the simplified method for accurate estimation of the C, F, P and W parameters. When possible risk estimates were generated, "brutal" combinations

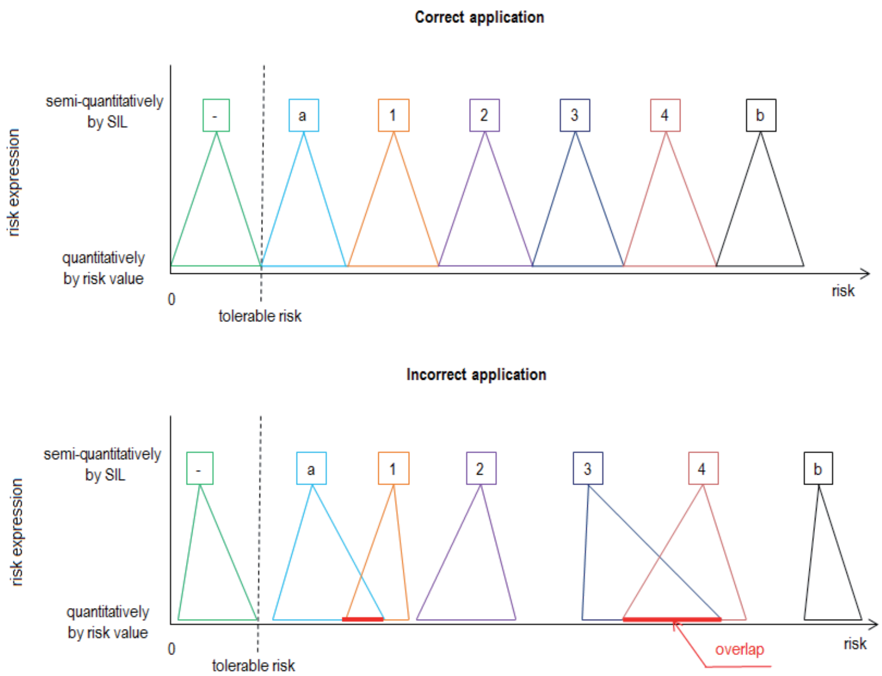


Fig. 5. Risk covered by SIL

Table 5. Number of overlaps (IEC 61508-5)

Overlap range	Number of couples of SIL category which overlap						
	0	1	2	3	4	5	6
One (next) SIL category	0	0	0	14	82	2 114	128 111
Two SIL categories	2	169	16 107	26 263	27 773	6 007	0
Three SIL categories	2 995	4 695	50 819	57 508	14 304	0	0
Four SIL categories	58 236	14 490	41 143	16 452	0	0	0
Five SIL categories	108 870	0	21 451	0	0	0	0
Six SIL categories	124 922	5 399	0	0	0	0	0

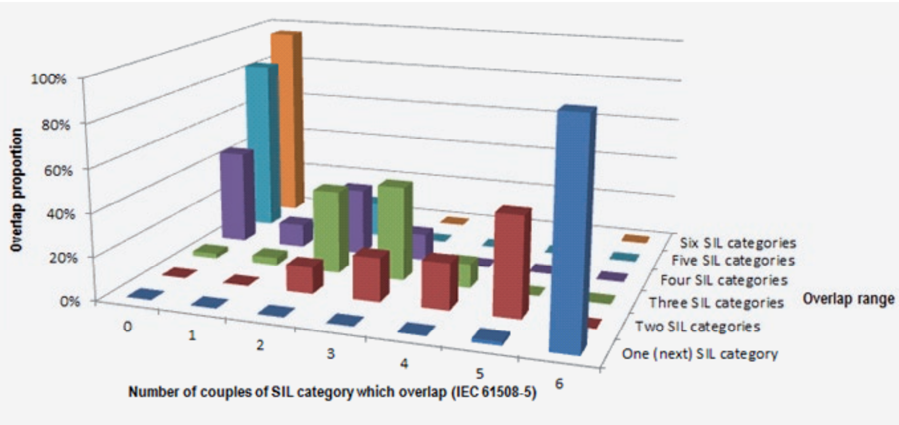


Fig. 6. Graph of overlap proportion (IEC 61508-5)

were also generated; these would not be used in a real application of the simplified method. These "brutal" combinations are represented for example by value 5 297 quotients combinations where the overlap goes over six SIL categories. Yet it is clear from the presented results that the simplified approach leads to significant inaccuracies in risk evaluation and SIL determination. Each combination of the *C*, *F*, *P* and *W* parameters determined SIL incorrectly, see value 0 in column "0" in Table 5. It cannot therefore be considered as correct.

3.3. The assessment of the correctness of the simplified approach according to IEC 62061

The method of determining safety integrity level (SIL) in this standard is different from the method specified in IEC 61585. It is based on semi-quantitative risk evaluation. It uses four parameters (*Se*, *Fr*, *Pr* and *Av*), which are assessed by points and reflected in the risk matrix where SIL determination takes place. Given that the score of probability parameters *Fr*, *Pr* and *Av* sums into a single endpoint parameter *Cl*, the use of geometric scales with the same quotient *Q* is apparent. Thus the sum arises from this relationship.

$$Q^{Fr} \cdot Q^{Pr} \cdot Q^{Av} = Q^{Fr+Pr+Av} \tag{5}$$

It represents the scales only on two levels, it is unnecessary to examine the scale composition (whether it is arithmetic or geometric) and can be considered as a scale drawn up in accordance with geometric sequence with an unknown quotient.

Again, as for IEC 6158-5, if the simplified approach is correct, the results must be consistent with the results obtained when using the fully quantitative risk evaluation. Each level of functional safety (1, 2, 3) covers successive risk intervals. When the correct parameters zones are set for *Se*, *Fr*, *Pr* and *Av*, they should not overlap. Therefore, inequality must apply.

$$R_1 < R_2 < R_3 \tag{6}$$

If the inequality is not satisfied, the risk is then overlapped by two or more SIL and the simplified approach cannot be considered as correct.

Results were found for all combinations of integers in the range <2; 20> quotients scales with geometric sequence parameters *Se*, *Fr*, *Pr* and *Av* by doing a simulation in Matlab. However, since parameters *Fr*, *Pr* and *Av* must always have the identical quotient, the possible number of quotients combinations is $19^2 = 361$. This is the total number of examined options for the simplified approach of risk evaluation. The observed number of overlaps is shown in Table 6 and Figure 7.

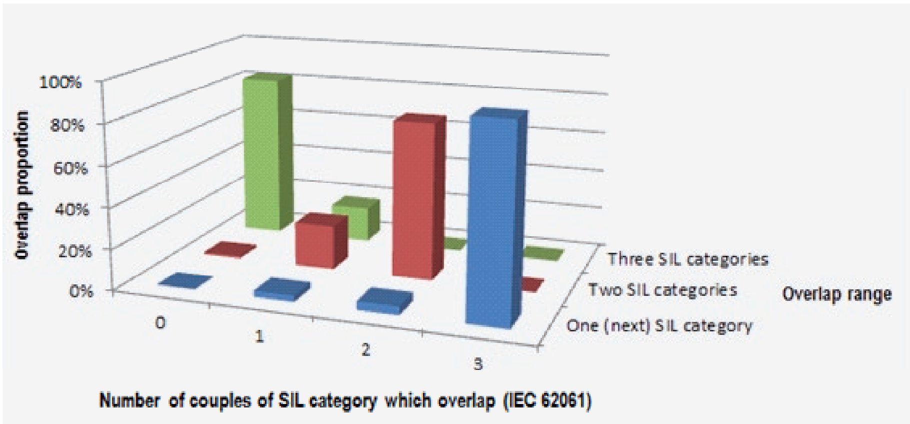


Fig. 7: Graph of overlap proportion (IEC 62061)

It can be clearly seen that even this simplified approach leads to significant inaccuracies in evaluating risk and determining SIL. Each combination of the *Se*, *Fr*, *Pr* and *Av* parameters determined SIL incorrectly, see value 0 in column "0" in Table 6. It cannot therefore be considered as correct.

Table 6. Number of overlaps (IEC 62061)

Overlap range	Number of couples of SIL category which overlap			
	0	1	2	3
One (next) SIL category	0	11	15	335
Two SIL categories	4	79	278	0
Three SIL categories	298	63	0	0

3.4. The assessment of the correctness of the simplified approach according to ISO 13849-1

Similarly to IEC 61508-5 the method for determining safety integrity level (PL) is in this standard is based on qualitative risk evaluation. It only uses three parameters. Parameter zones *S*, *F* and *P* and their scope are assigned with verbal description. Given that the scales only have two levels, it is then unnecessary to examine their composition (whether it is arithmetic or geometric) and can be regarded as scales based on geometric sequence with an unknown quotient. Again, as in IEC 61508-5, if the simplified approach is correct, the results must be consistent with the results obtained when using the fully quantitative risk assessment. Each level of functional safety (a, b, c, d, e) is covered by successive risk intervals. When the correct parameters zones are set for *S*, *F* and *P* they should not overlap. Therefore, inequality must apply.

$$R_a < R_b < R_c < R_d < R_e \qquad (7)$$

If the inequality is not satisfied, the risk is then overlapped by two or more SIL and the simplified approach cannot be considered as correct.

Results were obtained by doing a simulation in Matlab for all combinations of integers in the range <2; 20> quotients scales with geometric sequence parameters *S*, *F* and *P*. In total $19^3 = 6\,859$ possible risk evaluation options were examined according to the simplified approach. The observed number of overlaps is shown in Table 7 and Figure 8.

Table 7: Number of overlaps (ISO 13849-1)

Overlap range	Number of couples of PL category which overlap				
	0	1	2	3	4
One (next) PL category	2 109	0	4 750	0	0
Two PL categories	6 459	400	0	0	0
Three PL categories	6 859	0	0	0	0
Four PL categories	6 859	0	0	0	0

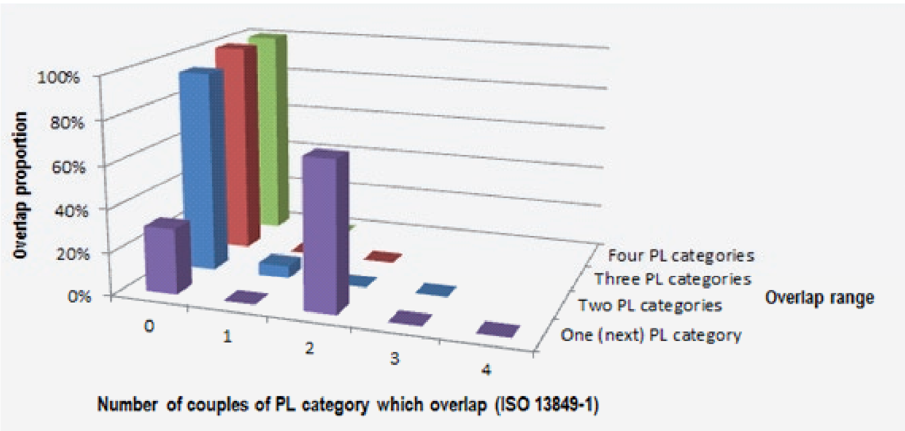


Fig. 8. Graph of overlap proportion (ISO 13849-1)

It can be clearly seen from the presented results that the simplified approach according to ISO 138491 is more robust than IEC 615085 and IEC 62061 in terms of resistance to inappropriate composition of risk parameter scales. This fact can be also observed from the concentration of values in column “0” in Table 7. Even this approach is not resistant against its improper use and it cannot be regarded as entirely correct.

4. Conclusion

IEC 61508 5, IEC 62061 and ISO 138491 standards do not indicate any primary sources with references to the fundamental works in the discipline around risk. The standards generally recommend applying quantitative risk evaluation but for practical application they only offer informative guidelines for qualitative and semi-quantitative risk evaluation in the form of graphs or risk matrices, from these then arise the requirements for SIL. Without taking into account the nature of the subjects that are sources of the risk, applying these standards can result into a series of serious errors.

In particular, simplified approaches of risk evaluation presented in the informative annexes of IEC 615085 and IEC 62061 are highly dependent on the correct understanding of the risks. They are very sensitive to the methods of composition scales of risk parameters. Therefore, the risk evaluation of complex systems (power plants, chemical plants, railway vehicles, etc.) should apply simplified methods uniformly to all devices that make up the entire system. Correct SIL determination cannot be guaranteed for qualitative or semi-quantitative risk evaluation of complex systems in accordance with those standards including ISO 13849-1. Fatal errors may occur; these can be partially eliminated by using a unified method for setting risk parameters (probability and consequence) for all suppliers of the equipment which then forms the complex system. This implies that the same range of scales should be used for all cases for assessing the probability and consequences. The risk of the equipment is in this case expressed implicitly. Tolerability of the risk is not clearly established. It is hidden. It is derived from the scales of probability and consequences; and the decision of which combination of probability and consequences begins to apply the lowest SIL.

This problem does not occur when the fully quantitative risk evaluation is used. Uniformity of the method used for evaluation is guaranteed. If the level of tolerability of risk is set the same for all devices of the complex system, there are no contradictions in determining the required SIL. The risk is expressed explicitly when the quantitative method is used. It is clearly defined whether the risk of the equipment is tolerable or not. This is due to setting the threshold value of tolerability for individual and societal risks, potentially economic or environmental risks. The requirements for SIL are clearly specified to achieve tolerable levels of risk.

***Acknowledgement**:*

This paper has been elaborated in the framework of the project Opportunity for young researchers, reg. no. CZ.1.07/2.3.00/30.0016, supported by Operational Programme Education for Competitiveness and co-financed by the European Social Fund and the state budget of the Czech Republic.

References

1. IEC 61508 5:2010, Functional safety of electrical/electronic/programmable electronic safety-related systems – Part 1: Examples of methods for the determination of safety integrity levels.

2. IEC 61511 x:2003, Functional safety – Safety instrumented system for the process industry sector.
3. IEC 62061:2005, Safety of machinery – Functional safety of safety-related electrical, electronic and programmable electronic control systems.
4. ISO 13849-1:2006, Safety of machinery – Safety-related parts of control system – Part 1: General principles for design.
5. IEC 61513:2001, Nuclear power plants – Instrumentation and control for systems important to safety – General requirements for systems.
6. EN 50129:2003, Railway applications – Communication, signalling and processing systems – Safety-related electronic systems for signalling.
7. http://www.automation.siemens.com/mcms/infocenter/dokumentencenter/ce/Documentsu20Brochures/e20001-a230-m103-v5-7600.pdf

Assoc. Prof. Pavel FUCHS
Technical University of Liberec
Studentska str., 2-46117 Liberec, Czech Republic
e-mail: pavel.fuchs@tul.cz

Jaroslav ZAJICEK, Ph.D.
Technical University of Ostrava
17. listopadu str.,15-70833 Ostrava – Poruba, Czech Republic
e-mail: jaroslav.zajicek@vsb.cz

Article citation info:
DĘBSKI H. Experimental investigation of post-buckling behavior of composite column with top-hat cross-section. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 106–110.

Hubert DĘBSKI

EXPERIMENTAL INVESTIGATION OF POST-BUCKLING BEHAVIOR OF COMPOSITE COLUMN WITH TOP-HAT CROSS-SECTION

EKSPERYMENTALNO-NUMERYCZNE BADANIA POKRYTYCZNYCH ZACHOWAŃ KOMPOZYTOWYCH KOLUMN O PRZEKROJU OMEGOWYM*

The object of this study is a thin-walled beam made of carbon-epoxy composite with open cross-section. The material used was a composite of epoxy matrix reinforced with carbon fiber (system HexPly M12, Hexcel). The M12 system is used above all in aircraft structures. It exhibits high fatigue durability and good maintenance properties at relatively low specific gravity. The research was lead as the FEM numerical analyses and experimental tests in buckling and post-buckling state, as well. In the conducted research in order to evaluate the effort ratio of the composite the Tsai-Wu tensor criterion was exploited. The numerical tool used was the ABAQUS software.

Keywords*: numerical modeling, thin-walled structures, FEM analysis, composite materials, experimental testing.*

Przedmiotem badań jest cienkościenna belka wykonana z kompozytu węglowo-epoksydowego o przekroju otwartym. Zastosowanym materiałem był kompozyt o osnowie żywicy epoksydowej wzmacniany włóknami węglowymi systemu HexPly M12 (Hexcel). System M12 wykorzystywany jest w przede wszystkim w strukturach lotniczych i charakteryzuje się wysoką trwałością zmęczeniową oraz dobrymi właściwościami eksploatacyjnymi, przy stosunkowo niskim ciężarze własnym. Badania prowadzono w zakresie obliczeń numerycznych z wykorzystaniem MES oraz badań eksperymentalnych w stanie krytycznym i pokrytycznym. W prowadzonych badaniach do oceny stopnia wytężenia kompozytu wykorzystano kryterium tensorowe Tsai-Wu. Zastosowanym narzędziem numerycznym był program ABAQUS.

Słowa kluczowe*: modelowanie numeryczne, cienkościenne struktury, analiza MES, materiały kompozytowe, badania eksperymentalne.*

1. Introduction

Thin-walled structures belong to a category of load-carrying structures that have wide field of applications in contemporary engineering. An example of their applications can be aircraft structures, for which high stiffness and strength demands under maintenance loads are formulated together with a tendency to minimize the mass of the structure. One of the basic questions connected to the design of thin-walled structures is a problem of stability loss, as well as load-carrying capacity of the system elements. Recently in the design of thin-walled load-carrying structures an increase in number of applications of modern constructional materials – composites in comparison to traditional structural materials could be noticed. Widely used group of materials are polymeric composites reinforced with glass, carbon or kevlar fibres. Application of these materials yields from advantageous ratio of their strength to mass and a resistance to unfavourable working conditions. A lot of papers concerning questions of stability, as well as load-carrying capacity of thin-walled structures is available [3, 11-14,16, 19, 25]. However, a large majority of them applies to classical structural materials, having isotropic properties. Over a span of a few recent years many articles describing properties of fibrous composites – the laminates were published, but they considered mainly theoretical models. There is still a lack of comprehensive information on experimental tests on layered composites reinforced with fibres. This inspired the author of this article to undertake a study in this subject area.

In this work, results of experimental tests on thin-walled composite columns of top-hat cross-section subjected to compressive load were presented. The obtained research results allow to verify the results given by FE models, as well as by author's own analytical-numerical (A-N) method based on the Koiter theory [10]. The experiments covered also determination of material properties, used later in worked out numerical models. Such approach allowed more credible comparison of the prepared composite profiles behaviour with numerical models, being usually only models of ideal structures.

2. Subject of research

The subject of the research were thin-walled columns of top-hat cross-sections made of M12/35%/UD134/AS7/300 Hexcel's "HexPly" unidirectional carbon-epoxy composite prepreg tape. Its matrix was made of epoxy resin (mass density: 1.24 g/cm^3; T_g: 128ºC; R_m: 64 MPa; ν: 0.4; E: 5.1 GPa), whereas the reinforcement was AS7J12K carbon fibers (mass density: 2.5 g/cm^3; R_m: 4830 MPa; ν: 0.269; E: 241 GPa). Nominal volume fraction of reinforcing fibres in the composite was ca 60 %. The composites were produced with autoclaving technique in the Department of Material Engineering at the Lublin University of Technology [5, 6, 15]. The laminate texture was composed of 8 plies of equal thickness of 0.131 mm sequenced symmetrically $[0,90]_{2s}$. The dimensions of the thin-walled column, as well as the composite layout are presented in Fig. 1.

The columns were produced with autoclaving technique with the use of vacuum packet, prepared in a special mould mapping the shape and the dimensions of the composite profiles. The prepared hermetic packet, providing stable sub-atmospheric pressure of ca -0.1 MPa was subjected to polymerization process in a laboratory-autoclave, where an overpressure of 0.4 MPa was kept in order to provide required holding down. In case of the carbon-epoxy composite a temperature

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

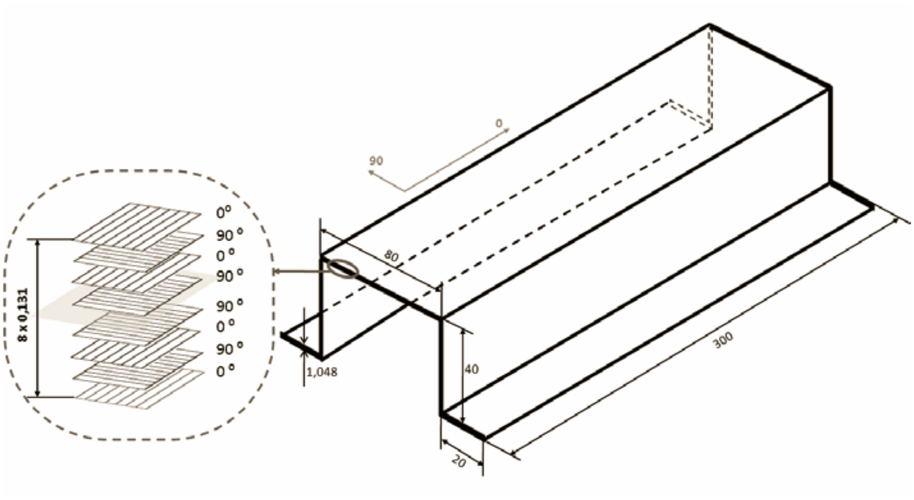


Fig.1. Dimensions of the analysed column and plies layup

body [7, 18]. The performed quality control of the produced profiles did not reveal any discontinuities in them. An uniform level of reflection of the entrance and of the echo from the bottom (A-scan), as well as uniform C-scan image was obtained. The B-scan was eliminated by decreasing the observation range down to values exceeding the B-scan usability (thin-walled elements) to the advantage of the C-scan module precision. The results of the measurements are given in Fig. 2.

Moreover, for the purpose of the laminate quality control microstructural testing with optical microscopy (NikonMA200, Japan) was employed. It based on computer image analysis (Image Pro Plus, NIS-Elements) and computer-assisted micro-tomography (SkyScan 1174 micro-tomograph). In particular, a quality of the profile's fillet radii was checked, as in these regions are especially prone to inter-layer discontinuities in the form of delamination. Texture inspection and non-destructive testing confirmed very good quality of the composite columns, especially in respect of material discontinuities (internal porosity, delamination). Application of autoclaving composite production technique enabled receiving structures having high mechanical characteristics confirmed by performed strength tests and minimal porosity amount <1%, as well as provided repeatability of the composite fabrication process.

of material heating of 135 °C was kept for 2 hours, what enabled finishing of the prepreg polymerization process. In order to eliminate disadvantageous phenomena usually emerging during the composite production process (excessive increase of thermal stresses inside the material and restraining of proper relaxation of initial and thermal stresses) a precise heating and cooling rate of 0.033 K/s was applied.

For the purpose of laminate texture quality check non-destructive testing (NDT) methods were used together with optical microscopy and computer-assisted micro-tomography. Each produced profile underwent thorough quality control in respect of flaw existence, such as porosity or delamination. Thus, every wall of the top-hat profile was examined with ultrasonic echo method using a phased array technique [9]. This testing was performed with OmniScan MXU-M ultrasonic defectoscope equipped with Olympus 5L64 A12 head and having wegde-type SA12-OL delay. The following test parameters were used: frequency 5 MHz, wave-propagation velocity 3100 m/s and amplification of 68 dB. In the conducted research for the purpose of quality control of the produced laminates an A-scan imaging (converted real-time amplitude plot), as well as a C-scan (a real-time collection of B-scan images coming from many converters). The imaging method used enabled simultaneous determination of a flaw's depth (A-scan), as well as its location depth and the flaw's width in any given direction (C-scan). The above mentioned procedure allows to localize delaminations or clusters of pores inside the composite

In order to determine mechanical properties of the produced laminates strength testing was performed in accordance with ISO standards. The experiments were done at room temperature (RT) with Zwick's Z100/SN3A universal testing machine, having the accuracy class of 1. To the prepared composite samples (columns) VISHAY's EA-13-24022-120 strain gauges were sticked. They were connected to the MGCplus (Hottinger) measurement system in order to measure the columns' deformations. A loading bar velocity was 2 mm/min. Experimentally determined basic mechanical characteristics of the carbon-epoxy composite were subsequently collected in Table 1. The received values were exploited in the definition of material model in Finite Elements numerical calculations.

Table 1. Mechanical characteristics of carbon-epoxy composite

Tensile strength F_{TU} [MPa]		Young modulus in tension E_T [GPa]		Poisson ratio ν		Shear strength F_{SU} [MPa]	Shear modulus G [GPa]	Compressive strength F_{CU} [MPa]	
0°	90°	0°	90°	0°	90°	±45°	±45°	0°	90°
1867.2	25.97	130.71	6.36	0.32	0.02	100.15	4.18	1531	214

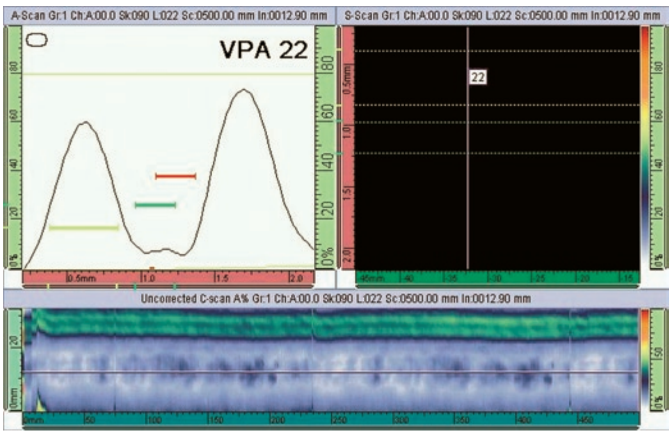


Fig. 2. Quality assessment of composite profiles with the use of non-destructive ultrasonic methods (NDT)

3. Numerical calculations

Discretisation of the thin-walled column model was based on 4-noded reduced integration shell elements (S4R) in the Abaqus and 8-noded elements (Shell99) in the Ansys software. Both types of finite elements had 6 degrees of freedom at each node. They were thin-walled plane-stress shell elements, in which the strain state was defined on the basis of displacements. On the other hand, the bending-state strains were described by angular displacements. A *Layup-Ply* technique was employed for the purpose of the symmetrical $[0,90,0,90]_s$ laminate modelling – Fig. 3.

Boundary conditions of the numerical model representing articulated support of the column's ends were defined by restraining the kinematic degrees of freedom of the nodes belonging to the borders of the first and the last cross-section. The load was applied to the model as uniformly distributed concentrated forces at the top-end of the column – Fig. 3.

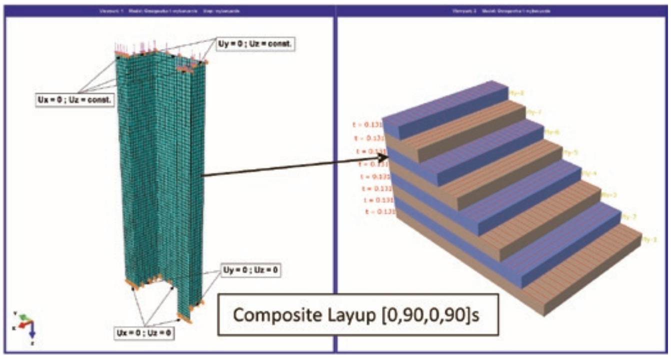


Fig.3. Discrete model of composite column

The properties of composite material were described by definition of orthotropic material in plane stress state, what allowed to describe the laminate properties in particular directions, according to fibres' arrangement [23] – Table 1. The numerical calculations within the framework of FE were performed in two stages. The linear stability problem (critical state) was solved by finding critical load and buckling mode, as well. Verification of the obtained results was performed with the A-N method [11], based on the Koiter's general asymptotic theory of conservative systems stability [10]. The post-critical calculations were performed as non-linear static analysis of the structure with initiated imperfections having a dimension of 0.05 of the top-hat profile wall thickness, corresponding to the first instability mode of the column. The post-critical analysis taking into account geometrically non-linear problem in Abaqus [1] was done with incremental-iterative Newton-Raphson method, whereas the post-critical equilibrium path was tracked with the Arc-length procedure (the Ricks Method) in the Ansys program [2]. The range of the numerical calculations covered also an attempt to estimate a probability of the composite damage occurence in post-critical state according to the Tsai-Wu criterion [26]. This needed determination of additional material characteristics, such as: F_{TU} – tensile strength along fibres (0º-direction) and in the perpendicular 90º-direction, as well, F_{CU} – compressive strength in both directions and F_{SU} – shear strength of the ±45º-interface. These characteristics were obtained from the performed experiments, see Table 1.

4. Experiments

Stand tests of the top-hat cross-section columns in compression were done with the Zwick Z100/SN3A 1-accuracy class universal testing machine of the 100 kN load range. For this purpose the machine was equipped with intentionally designed and fabricated grips providing that a column was loaded axially – Fig. 4. The grips were set coaxially by fixing them to the machine's loading pins. The ball-and-socket joint enabled free rotations of the grips. Small imperfections of the composite columns ends, as well as possibility to occur local effects in boundary sections of the columns were compensated with thin soft-plastic pads. Before each test the loading system was loaded up to 15 % of the expected critical load in order to provide best alignment of the column placed between the grips. Next, the grip retainers were removed and the column was completely unloaded. On the sample's surface, in a point of the biggest deflection of the composite profile web (the middle wall) two strain gauges were sticked along the 0º-direction on both sides of the column. In addition, the deflection was measured with the OptoNCDT 1605 laser sensor at the point of the biggest deflections of the profile web or its arm (the side wall).

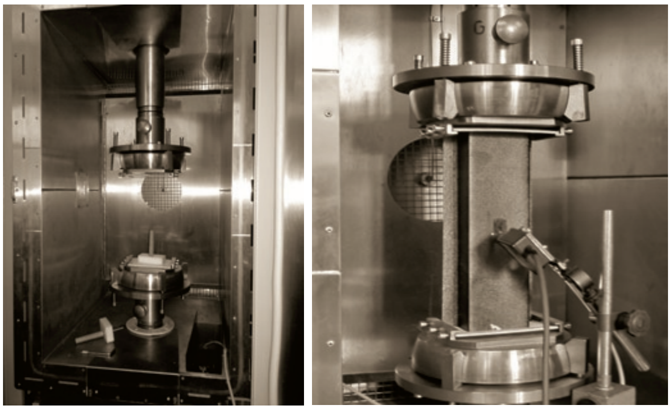


Fig.4. Test stand outfit: self-aligning grips, composite column and laser sensor

During the tests the following variables were registered: the time, the compressive force, the displacement of the cross-bar, the web deflection (with the laser sensor) and the strains (with gauges). The sampling frequency of all parameters was 1 Hz. The experiments were lead in standard conditions at 23 ºC and at a steady cross-bar velocity equal 1 mm/min. The tests were continued until the load reached double critical force. The tests covered registration of the subcritical, critical and post-critical state. However, during the tests no symptoms of the structure failure were perceived.

In the conducted experiments for the purpose of critical forces' estimation the following methods were used [8, 20, 21, 22, 24, 27, 28]:

a) the vertical-tangent line method (the mean-strains method) – denoted as K1,
b) the method of straight lines intersection in the plot of mean strains – denoted as K2,
c) the P-w^2 method – denoted as K3,
d) the inflexion-point method – denoted as K4,
e) the Tereszowski method – denoted as K5,
f) the Koiter method – denoted as K6.

The experimental tests were conducted on 3 specimens with 3 measurements for each of them.

5. Results

The critical state analysis of the compressed thin-walled column showed a local mode of stability loss, manifesting itself by taking a shape of 4 half-waves by all walls of the profile – Fig. 5. For every numerical tool employed in simulations (FEM, A-N method) a qualitative, as well as a quantitative agreement of the computational results was obtained.

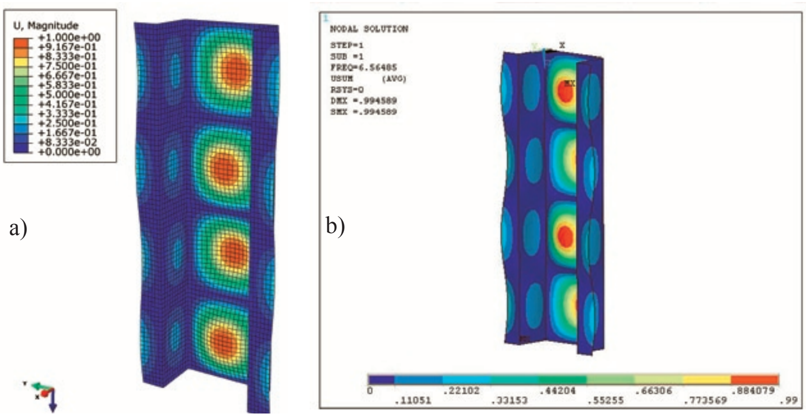


Fig.5. First buckling mode of composite column: a) Abaqus results; b) Ansys results

Table 2. Critical load values for the firs buckling mode

Computational tool	Abaqus (FEM)	Ansys (FEM)	Analytical-numerical method (A-N)
Critical force [N]	$P_{CR/Ab} = 6655$ [N]	$P_{CR/As} = 6565$ [N]	$P_{CR} = 6629$ [N]
Number of half-waves	4	4	4

The values of critical forces obtained with particular methods were collected in Table 2. Experimental results for the critical state were given below in a form of force vs testing method plots, obtained from all measurements performed with any particular method, together with confidence interval – Fig. 6. Non-linear computations enabled deformation mode analysis of the structure in post-critical state, up to a laminate's failure load, determined with the Tsai-Wu criterion. The failure load value for the laminate material was accepted as the one corresponding to a failure parameter equal to 1 (on the scale of 0 to 1). The failure force values determined this way with the Abaqus and the Ansys programs were $P_{f/Ab} = 18.6$ kN and $P_{f/An} = 17.1$ kN, respectively. One can notice, that these values equal to 280% or 260% of the respective critical force $P_{CR/Ab}$ or $P_{CR/An}$. Zones in models, in which critical value of the failure parameter was reached indicate hazardous regions of the real composite structure. This means, the regions where the risk of damage of some plies is high. The post-critical deformation form with Tsai-Wu criterion maps for the upper surface of the external laminate layer obtained with the Abaqus and the Ansys programs at failure load are shown in Fig. 7.

Figure 8 presents a comparison of post-critical force-displacement equilibrium paths for the node experiencing maximal deformation amplitude of the wider wall of the top-hat profile. The plot displays the results obtained with different research methods: FEM (Abaqus), analytical-numerical (A-N) method, as well as experimental outcomes.

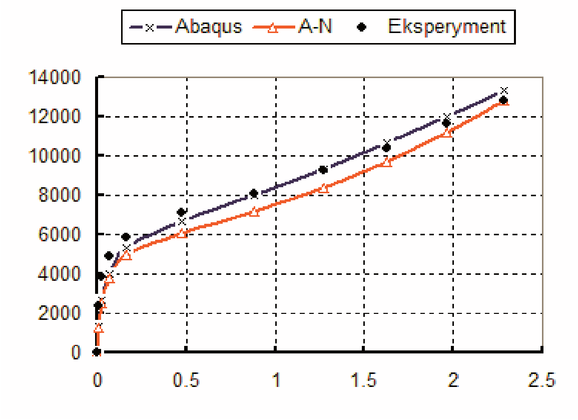
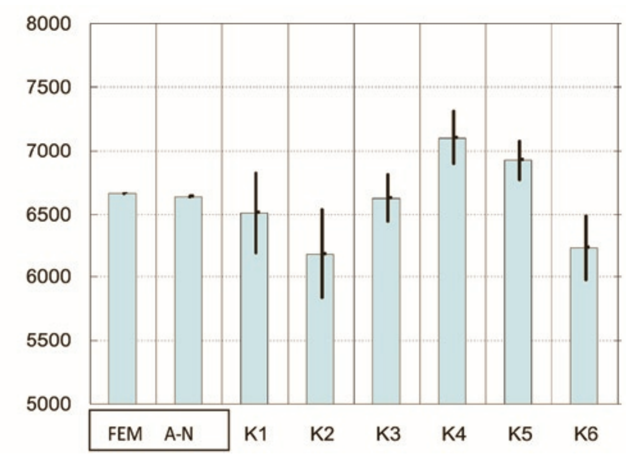
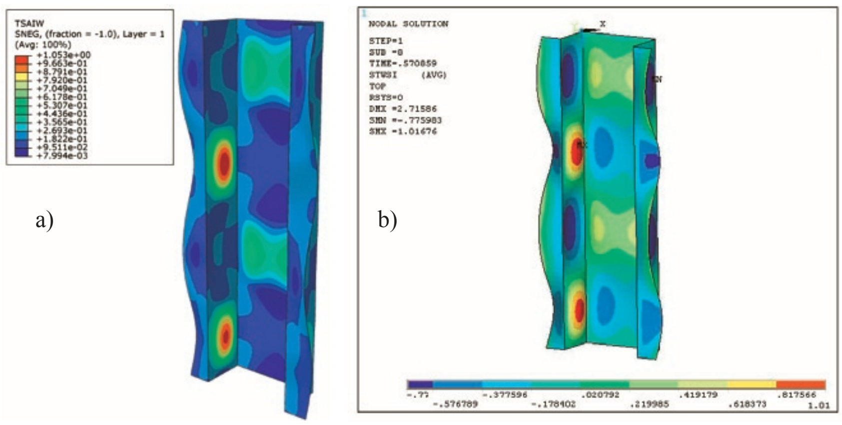


Fig.8. Post-critical force-displacement equilibrium paths



Method	P_{CR} [N]
FEM/Ab	6655.50
A-N	6629.10
K1	6503.85
K2	6184.84
K3	6619.88
K4	7102.08
K5	6918.98
K6	6227.82

Fig.6. Comparison of critical forces' values PCR [N] determined with different methods



a) b)

Fig.7. Post-critical deformation: a) Tsai-Wu – Abaqus; b) Tsai-Wu – Ansys

6. Conclusions

The paper presents studies on critical, as well as post-critical state of thin-walled composite columns subjected to axial compression. The performed analysis prooved qualitative and quantitative agreement of the research results made with different methods. Analysis of the curves shown in Fig. 8 reveals good agreement of computational results with those of experiments, both in subcritical and post-critical range. This confirms the adequacy of the worked out numerical models. The accepted methods of critical load determination, based on experiments (Fig. 6) allowed to estimate the range of experimental value of the critical force $P_{CR} = 6184 \div 7102$ N with maximal difference among the used methods equal to 13 %. The critical load value determined with computational methods: FEM (the Abaqus, the Ansys) and analytical-numerical (A-N) method were located in the middle of the obtained range – Table 2. In the post-critical range almost identical force-displacement equilibrium paths in case of FEM and experiment were obtained. Only the A-N method gave slightly different results (ca 10 %), but the nature of the curve was the same.

The obtained results give wide possibilities of observation and analysis of deformation states, as well as effort levels up to the failure. This enabled the identification of damage-prone zones in the laminate and the determination of the failure load level in relation to the critical load. The analysis of the post-cricital equilibrium path allows to assess the structure's stiffness after the loss of stability in the context of applied ply sequence. Thus, the obtained results delivered a lot of important information useful in the process of forming and optimization of the composite texture in the context of its maintenance loading conditions.

Acknowledgements: *An article written under the ministerial research project no. N N507 241440 The Ministry of Science and Higher Education*

References

1. Abaqus HTML Documentation.
2. Ansys HTML Documentation.
3. Bazant ZP, Cedolin L. Stability of structures. Elastic, inelastic, fracture and damage theories. Oxford University Press 1991.
4. Bieniaś J, Dębski H, Surowska B, Sadowski T. Analysis of microstructure damage in carbon/epoxy composites using FEM, Computational Materials Science, 2012, 64:168–172.
5. Campbell FC, Manufacturing Processes for Advanced Composites, Elsevier Ltd, 2004.
6. Campbell FC, Manufacturing Technology for Aerospace Structural Materials, Elsevier 2006.
7. Capitani V, Capriolo M, Sendi D, Characterization of Casting Defects in Composite Carbon Fiber Material Detected by Ultrasonic Inspection, 2011.
8. Coan JM, Large-Deflection Theory for Plates With Small Initial Curvature Loaded in Edge Compression, ASME, Journal of Applied Mechanics, June 1951, 18:143–151.
9. Habermehl J, Lamarre A, Ultrasonic Phased Array tools for composite inspection during maintenance and manufacturing, 17th World Conference on Nondestructive Testing, Shanghai, China, 2008.
10. Koiter WT. Elastic stability and post-buckling behavior. In: Proceedings of the Symposium on Non-linear Problems. Univ. of Wisconsin Press, Wisconsin, 1963, 257–275.
11. Kolakowski Z, Kowal-Michalska K (Eds.). Selected problems of instabilities in composite structures, A Series of Monographs, Technical University of Lodz, 1999.
12. Kolakowski Z, Kubiak T. Load-carrying capacity of thin-walled composite structures. Composite Structures 2005; 67 (4): 417–426.
13. Kowal-Michalska K (ed.). Dynamic Stability of Composite Plate Structures, (in Polish), WNT, Warszawa, 2007.
14. Królak M. and Mania RJ (eds.). Statics, dynamics and stability of structures. Vol. 1. Stability of thin-walled plate structures, Technical University of Lodz, Series of monographs, Lodz 2011.
15. Mason K. Autoclave Quality Outside The Autoclave?, High Performance Composites 3/1/2006.
16. Niezgodziński T, Kubiak T. The problem of stability of web sheets in box-girders of overhead cranes. Thin-Walled Structures 2005; 43 (12): 1913–1925.
17. Pałka K, Bieniaś J, Dębski H, Niewczas A. Funite element analysis of thermo-mechanical loaded teeth. Computational Materials Science 2012; 64: 289–294.
18. Schnars U, Henrich R. Applications of NDT Methods on Composite Structures in Aerospace Industry, Conference on Damage in Composite Materials, Stuttgart, Germany 2006.
19. Simitses GJ, Hodges DH. Fundamentals of structural stability. Butterworth-Heinemann 2006.
20. Singer J, Arbocz J, Weller T. Buckling Experiments. Experimental methods in buckling of thin-walled structure. Basic concepts, columns, beams, and plates, Volume 1, John Wiley & Sons Inc. New York 1998.
21. Singer J, Arbocz J, Weller T, Buckling Experiments. Experimental methods in buckling of thin-walled structure. Shells built-up structures, composites and additional topics, Volume 2, John Wiley & Sons Inc. New York 2002.
22. Spencer HH. Walker AC. Technique for Measuring The Critical Loads of Columns and Plates, Paper No. 2290 A SESA Spring Meeting, 1974.
23. Swanson SR. Introduction to Design and Analysis with Advanced Composite Materials, Prentice-Hall, Inc., 1997.
24. Tereszkowski Z, An experimental method fordetermining critical loads of plates. Archive of mechanical engineering 1970; 3: 485–493.
25. Thompson JMT, Hunt GW. General theory of elastic stability. Wiley, New York 1973.
26. Tsai SW. Introduction to Composite Materials. Technomic 1980.
27. Van der Heijden AMA (red.). W.T. Koiter's Elastic Stability of Solids and Structures. Cambridge University Press, 2009.
28. Venkataramaiah KR, Roorda J. Analysis of local plate buckling experimental data, Sixth international specialty conference on cold-formed steel structures (1982: November 16–17; St. Louis, Missouri). Missouri S&T (formerly the University of Missouri-Rolla), 1982, 45–74.

Hubert DĘBSKI, Ph.D. (Eng.)
Department of Machine Design
Lublin University of Technology
ul.Nadbystrzycka 36, 20-816 Lublin, Poland
e-mail: h.debski@pollub.pl

Article citation info:
KAMARUZZAMAN S-N, MYEDA NE, PITT M. Performance levels of high-rise private office buildings maintenance management in Malaysia. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 111–116.

Syahrul-Nizam KAMARUZZAMAN
Nik Elyna MYEDA
Michael PITT

PERFORMANCE LEVELS OF HIGH-RISE PRIVATE OFFICE BUILDINGS MAINTENANCE MANAGEMENT IN MALAYSIA

POZIOMY WYDAJNOŚCI ZARZĄDZANIA UTRZYMANIEM RUCHU W PRYWATNYCH WIELOKONDYGNACYJNYCH BUDYNKACH BIUROWYCH W MALEZJI

Maintenance management is an essential aspect in determining the performance and quality of properties such as office buildings. The fundamental issues related to techniques and approaches concerned are generally taken lightly by the practitioners which lead to inefficiency of maintenance management practice in the market today. The paper aims to determine the current standard and performance of maintenance management system by applying the study to high-rise private office buildings. These objectives are to be achieved by evaluating and analysing perceptions of the end users from five (5) high-rise office buildings in Klang Valley by using a mixed method combination of both quantitative and qualitative methods. Research findings signify that the performance of high-rise office buildings are generally rated as average by the end users and results from interviews with the maintenance managers denote the detailed systems applied. There is a positive relationship between the maintenance management systems and performance of maintenance management. This paper provides an important research which uncovered the scenario in the industry and the key perceptions by the building end users. This research is anticipated to be significantly beneficial and can be further used as a piece of information specifying on high rise private office buildings.

Keywords: *Maintenance management, performance levels, end users, high-rise office buildings, performance measurement.*

Zarządzanie utrzymaniem ruchu jest ważnym aspektem przy określaniu wydajności i jakości mienia, takiego jak na przykład budynki biurowe. Podstawowe zagadnienia związane z odpowiednimi technikami i metodami są na ogół traktowane dość pochopnie przez praktyków, co może prowadzić do powszechnej nieefektywności zarządzania utrzymaniem ruchu. Celem artykułu jest określenie bieżącego poziomu i wydajności systemu zarządzania utrzymaniem ruchu poprzez badania, których przedmiotem są wielokondygnacyjne prywatne budynki biurowe. Cele te zrealizowano poprzez ocenę i analizę obserwacji użytkowników końcowych z pięciu wielokondygnacyjnych budynków biurowych w Dolinie Klang przy użyciu metody mieszanej łączącej podejście ilościowe i jakościowe. Wyniki badań pokazują, iż wydajność wielokondygnacyjnych budynków biurowych jest ogólnie oceniana przez użytkowników końcowych jako przeciętna, podczas gdy rezultaty wywiadów ze specjalistami utrzymania ruchu dostarczyły informacji nt. stosowanych systemów. Istnieje pozytywny związek między systemami zarządzania utrzymaniem ruchu oraz jego wydajnością. Niniejszy artykuł przedstawia ważne badania, które prezentują praktyki stosowane w przemyśle jak również kluczowe spostrzeżenia użytkowników końcowych budynków. Badania mogą mieć korzystne przełożenie; mogą także znaleźć dalsze zastosowanie jako źródło szczegółowych informacji nt. wielokondygnacyjnych prywatnych budynków biurowych.

Słowa kluczowe: *zarządzanie utrzymaniem ruchu, poziomy wydajności, odbiorcy końcowi, wielokondygnacyjne budynki biurowe, pomiar wydajności.*

1. Introduction

Maintenance Management or Operations Management has been defined as a function that transforms input including people, capital, energy, materials and technology into outputs namely goods and services [22]. Coetzee [7] defines it as an activity that aims to optimise the availability and reliability of production equipment and maintain its operability at an acceptable cost level. Subsequently, Certo [6] has labeled operations management as a systematic direction and control of operation processes. These definitions have highlighted maintenance management as a systematic design used by the people that operate the organizations to control the overall operation processes in transforming the inputs into goods and services.

Bernard et al. [5] report that the 'Deferred Maintenance Concept' that was well known 10 years earlier was still ongoing in the maintenance department. The deferred maintenance concept that is basically of postponing works and accumulating overdue maintenance works has entailed enormous amounts of money to be expanded. He also claimed that facilities were aging and constantly being renovated in piecemeal fashion which means that the actions were taken at different times or ways rather than carefully planned from the beginning. Bernard et al. [5] also add that many renovations were limited in scope due to funding restraints and many times result in cosmetic change with few or no infrastructure improvements. This fragmentation could further lead to inefficiency of mechanical systems, customer complaints due to dissatisfaction with their facility or space environmental conditions, and eventually higher utility bills and maintenance costs.

Meanwhile, Hinks [11] relates a maintenance management performance scenario from his interviews with the facility managers to find their agreed set of indicators. The managers were uninterested in considering any facets of maintenance management performance below an aggregated level of indicator for maintenance. He also claims that the business managers did not consider any management details as they relied more on reactive actions based on clients' or users' complaints.

This indicates that maintenance managements are still being practiced in improper procedure by the maintenance managers which subsequently caused bad impacts to the facilities and the services provided. It can be seen that the managers prefer carrying out reactive maintenance works rather than proactive works and at times do not consider for clients satisfaction and also the performance of services. There is an increasing concern that the maintenance management has been unprofessionally applied by the maintenance managers and no research has so far outlined the critical factors and deliberation on such impractical practices.

While Gelders et al. [10] suggested that there are four (4) measurements of benchmarking to be looked at like financial (shareholder's views), customers, internal processes (the long- and short-term means to achieve financial and customer objectives), learning and growth (capability to improve and create value). In this scope, customer group that comprises of either clients or building occupants can be also known as end users. It is undeniably true that end users' perceptions and satisfaction level are able to conduct the maintenance managers to choose the right channel and implementation to upgrade the building performance.

Therefore, maintenance management is conclusively proven as an essential facet of a property's performance. An efficient maintenance management will produce a systematic and excellent maintenance management which increases the operation productivity and performance whereas improper conduct decreases the performance level and affect the life cycle of a property. Evaluation on the end users' response can lead to the building performance analysis which will later help to identify the gaps existed between the service provided and satisfaction level. Therefore, measurement is needed to be evaluated in order to maintain an impressive standard of performance and benchmarking is suggested as one highly influential tool to identify actual and current performance of a practice to suit the best practices available in the market.

2. Maintenance Management: Roles & Current Scenario

In emphasising the importance of maintenance management in the property industry today, various literature concerning maintenance management and end users' perceptions and satisfaction key factors are reviewed. The significance of maintenance and its position in the world's diverse industries can be seen in progressive developments of manufacturing, refineries, mining and building. Egbu [9] signifies the role of maintenance as the major driver of economic growth whereby it generates 45-60% of fixed capital formation in many countries and also generates 5-15% of Gross Domestic Product (GDP). The role of maintenance in modern manufacturing is becoming ever more important with companies adopting maintenance as a profit-generating business element [14].

While in refineries, the maintenance and operations department are very large and each department consists of up to 30% of the total staffing [8]. A study by the Swedish mining industry shows that the cost of maintenance in a highly mechanised mine can be 40-60% of the operating cost [7]. Facilities and maintenance management also contributes from 5 to 10% of employment in individual countries whereby it supplies approximately 111 million people which constitutes the majority of labour force which is 75% in developing countries [9].The role is particularly effective in developing countries due to the rapid and large-scale urbanization which require large scale of facilities and maintenance management.

Wordsworth [25] reports that building maintenance accounts for over half the building industry's total output, and for over two third of the contracts let. Subsequently, the role of maintenance manager continues to expand, as more demands are made by users regarding the economic and functional efficiency of the buildings in which they live and work. Maintenance provides critical support for heavy and capital-intensive industry by keeping machinery and equipment in a safe operating condition [21].

Tsang [24] opines that maintenance works as an important support function in business with significant investment in physical assets and plays an important role in achieving organizational goals. Ahmad [1] points that property for instance building is a valuable physical asset that requires an efficient maintenance management. Yasmin [26] also believes that any organisation which manages high-rise commercial buildings must emphasise the effectiveness and efficiency in their operations. This can also be related to the office buildings as it also requires a good maintenance management system. The condition and quality of the buildings also reflect public pride or indifference, the level of prosperity in the area, social values and behaviour and all the many influences both past and present combine to give a community its unique character.

Nevertheless, improper conduct and application of maintenance management procedure and systems may result in deteriorating the property itself. The impact may be seen in demoting a planned financial costing and the loss in value of the property. As such, an excellent practice of maintenance management is needed to increase the life cycle of property and to minimise unexpected breakdowns or deterioration effects. Therefore, the performances of maintenance management operations have to be continuously reviewed and analysed in order to ascertain a high quality service.

According to Egbu [9], buildings in overall contribute 33% to CO_2 emissions, which gives a substantial impact to the environment.

Table 1. Chronology of the Building Defects Occurrence.

Year	Chronology Of The Occurrence Of Building Defects
2005	Collapsed ceiling at the Parliament House, Jalan Duta Kuala Lumpur
2006	Fungus defects on wall at the Hospital Sultanah Aminah, Johor Bahru
2007	Defects at the Navy Recruit Training Centre (PULAREK), Johor
2007	Floods from 7th Floor down to 2nd Floor of the Immigration Department, Putrajaya
2007	Collapse of plaster ceiling at the Entrepreneur and Co-operative Development Ministry, Putrajaya
2007	Collapse of ceiling at the new court complex at Jalan Duta, Kuala Lumpur
2007	Collapse of ceiling at the Parliament House,Jln Duta, Kuala Lumpur
2007	Floods caused by leaking pipes, roof and wiring problems at the new court complex at Jalan Duta, Kuala Lumpur
2007	Collapse of ceiling at the Hospital Sultan Abdul Halim, Sg. Petani, Kedah

Legislation and stakeholder concern increasingly require facility managers to reduce CO_2 emissions. In this respect, the management of buildings needs to be emphasized and systematically controlled. The government and office buildings managers need to take cognizance of this and plan for a better eco-friendly management. In 2006, the government has allocated about one trillion Ringgit towards maintaining the public building facilities [9]. However, in the Malaysian context, the government is yet to implement any guidelines for maintenance management and also the performance measurement of the system applied [3]. Therefore, maintenance agents or companies from both public and private sectors have no systematic guidelines to be followed and no specific compliance to be adhered to in order to deliver for the best [20].

3. Performance Measurement

Several frameworks have been developed for measuring performance over the years. Until 1980, the performance measurement was based on mostly financial measures [23]. According to Kaplan and Norton [13], the approach at that time looks into 4 perspectives that focus on financial aspects, customers, internal processes and innovation and learning. Subsequently, various researchers have developed frameworks considering non-financial measurements and intangible assets to achieve competitive advantages [13].

A performance measurement system is developed by author for the research methodology purposes by incorporating the common maintenance management systems applied into it (Figure 1). Based on the literature review, the performance or maintenance indicators are identified complete with the performance indicators respectively. Functional indicator for instance outlines the management service delivery as its significant aspect. In this scope, the research assesses the performance of the service based on the characteristics such as reliability in which the assurance or confidence delivered by the managers, responsiveness as to whether positive or negative response is given and also timeliness which emphasizes on the promptness of response or action taken.

As for the technical indicator, building maintenance with detailed list of the maintenance dimensions are identified with reference to the literature review and basic services provided generally by office building managers for instance cleaning, landscaping, general maintenance, lightings, air-conditioning, lift or escalators, mechanical and electrical (M&E), sanitary & plumbing, access, signage and also parking. These maintenance services are regarded as the backbone of maintenance management of an office building. This is in accordance with opinion from Egbu [9] which explains that the importance of building services to the success of an organization has never been greater and continues to grow. Alternatively, for image indicator, focus on the quality of external and internal finishes of the building.

The three maintenance indicators namely functional-management service, technical-building maintenance and image-building image with respective dimensions are to be measured with both focus groups that are maintenance managers and end users. Different performance key factors are designed for the focus groups like time, quality and costs are targeted as the benchmarking or key factor to measure the level of performance for maintenance managers' scope while only time and quality factors are designated for end users as cost factor is most likely be unsuitable to be measured for this group.

All elements in this system are in overall interrelated and play important roles in sustaining the overall performance of maintenance management. These elements also meet the characteristics defined by Al-Sultan and Duffuaa [2] as it is believed to be relevant, interpretable, valid, time effectiveness and also easy to be implemented.

4. Significance of Customers' Opinions

A fundamental premise to the service concept is the notion of satisfying the customers' needs. A satisfied customer is to enhance a service and the firm's bottom line in multiple ways. Increased customer satisfaction generates positive reviews and brings in new customers to the firm. There are many studies that highlight the importance of customer satisfaction for a firm's success and how customer satisfaction can be measured. However, most of these studies are limited to the area of business-to-consumer marketing and not for the maintenance services provided. Myeda et al. [19] also signifies that maintenance managers should value the important roles of end users in evaluating the performance of maintenance services with a great attention given on their needs and requirements. Consultations with the end users should be a mechanism to establish a proactive management process.

Therefore this study determines to explore customer satisfaction as a function of end users' perceptions about the maintenance services

Maintenance Aspects	Performance Measurement Dimensions	Focus Groups	Performance Key Metrics
Functional Management Service Delivery	**Delivery Characteristics** *Assurance* *Reliability* *Responsiveness* *Timeliness* *Validity*	• Maintenance Managers • End Users	• Time/ Use • Quality
Technical Maintenance Services	**Maintenance Services** *Cleaning & landscaping* *General Maintenance* *Lightings* *Air-Conditionings* *Lifts/ Escalators* *M&E* *Sanitary & Washing Facilities* *Access, Signage & Parking* *Safety & Security*		
Image Building Image	**Building Image** *External Image* *Internal Image*		

Fig. 1. Performance Measurement System designed

and then relate the importance of these perceptions to the performance level of the maintenance services provided. This is as customers or in this research more accurately referred as end users gives the most accurate results and perceptions of the services delivered to them. According to Spires [23], there is a clear trend towards customers demanding industry specialized systems. The pressures and influences on suppliers to accommodate this demand are vast, but unless a supplier is of sufficient size to afford the continual improvement required for product development, they will struggle to achieve profitability and long-term stability becomes less likely.

As a result, smaller companies will either become niche specialists, or within the asset and maintenance management market. Form strategic alliances with one or more of the bigger companies, or quite simply go out of business. However, even for large companies, such development costs are a major expense, and to supply a unique system for each customer is not practicable. The solution is products that are modular, off the shelf, but also highly configurable to suit each customer's requirements [23]. Through customer demands we will see a rationalization of suppliers of such solutions, and companies operating in this sector today may not necessarily survive in their existing form tomorrow. Product offerings will also change and evolve into modular, industry-standard systems, which through extreme flexibility will allow enough bespoke capability to satisfy individual customer requirements, while techniques such as Realiability Centered Maintenance (RCM) will gain acceptance as more firms strive to achieve *"best practice"* [25].

Finally, there are also wider complications for the whole of industry. According to Spires [23], existing evidence has identified that the most successful implementations of asset and maintenance management solutions are in companies, which embrace the concept of "*best practice*" as a total company culture. As a result, more and more companies will look to establish overall "*best business practice*" which indirectly benefits industry in general.

5. Research Methodology

Pragmatic Worldview

This study is based on pragmatic principle that is suitable for mixed method research, where Morgan [18] signifies that the approach is to focus on the research problem in social science and then using pluralistic approaches to derive knowledge about the problem. Johnson and Onwuegbuzie [12] also believes that pragmatic principle also justify for the methodology of combining both qualitative and quantitative approaches. This is supported by Mertens [16] in which studies in communities using mixed methods approach have the potential to address persistent inequalities and challenging social conditions. This can be achieved by an exploration of both the study's strength and challenges through multiple methods. Mixed method is also a systematic inquiry into the variations of social constructions of meaning among interview and survey respondents may not only help in validating research instruments and scales, but may go further in that they could produce complementary subsets of results, which would enrich overall study [4]. This also helps in shaping better exploration, guide analysis and interpretation.

Mixed Method Design

Leech and Onwuegbuzie [15] also suggested that there are three dimensions of mixed method design. The typology is conceptualized on the level of mixing (partially mixed or fully mixed), time orientation (concurrent or sequential) and emphasise on approaches (equal status or dominant). As for this study, both quantitative and qualitative data were connected throughout the data collection and analysis phases. The conclusions of the study were drawn between the two bases of data analysis and findings. This study adopts the fully mixed methods, which is also the highest degree of mixing research methods and research paradigm characteristics [15]. This study fits the criteria as it adopts both quantitative and qualitative approach across the other components namely the research objective, type of data and operation methods, type of analysis and finally type of interference). The term fully mixed also refers to the fact that the mixed methods approach was implemented at various stages of the research including in the data analysis process where some of the qualitative data were converted into numbers. There are also various preceding studies that have converted qualitative data into numbers [17].

The time orientation is concurrent, where both questionnaires survey and case study interviews being conducted at the same time. The weightage is also the same where the status is equal as the end results are equally important to achieve the objective of the study.

Data Collection

The research adopted mixed method; a combination of both quantitative and qualitative methods. Five (5) high-rise office buildings in the vicinity of Kuala Lumpur were selected for the case studies whereby the important focus groups involved were maintenance managers and respective end users. Multiple data collection techniques were used in qualitative and quantitative data collection processes including focused interviews, case studies and observation. Evidently, the case studies data collection technique is appropriate to achieve the approach of our study that is to study and analyse the management systems and end users' perceptions from five (5) different office buildings.

Interviews were conducted specifically to the maintenance managers of five (5) chosen high-rise office buildings. Maintenance managers were enquired on information pertaining the building background, maintenance services provided, systems used, manpower, subcontractors as well as problems and improvements that have been completed or in progress for the building. The questions were prepared based on the semi-structured and also open-ended ad hoc conversations. Results of the interviews were necessary for the research evaluation on the systems applied by the maintenance managers and subsequently for the relationships identification between the main two variables.

Observation is also one vital step to fulfill the objective of this research as it will record the patterns of certain scenarios or behaviours occur in specific settings. Observations were conducted during site visit to the chosen case study areas and also while conducting the interviews. It was based on direct observation on the processes involved of the maintenance management system at each of the case study premises.

Questionnaire survey captures information through the input of responses to a research instrument containing questions that was through the 252 sets of questionnaires distributed to the five (5) office buildings respectively. Researcher has personally distributed the questionnaires to the respondents and opted for follow up phone surveys for the late respondents. Questionnaires were distributed from June 2008 and collected in late August for data analysis. Sets of questionnaires with structured and semi-structured / open-ended questions were distributed to the respective building end users so as to discover regularities among groups of maintenance management by comparison of answers to the same set of questions. The analysis of data from the questionnaires responses provides precise data from which tables and graphs are produced.

Data Analysis

The steps involved in analysing the data started with categorising all the results and also putting them into appropriate groups for coding purposes.

Results for the maintenance management system were derived from the checklist on their conformity to the respective maintenance management system maintenance aspects; functional, technical and building image. The characteristics of service conformity for the maintenance aspects were drawn from literature review and categorised as shown in Table 2.

The Statistical Analysis Software 16.0 was used to analyse the correlation between the performance of maintenance management and also maintenance management systems applied. The correlation coefficients will indicate the correlation strength, which will also highlights the impact that each factor represents.

6. Discussions & Findings: Relationship between Maintenance Management System and Performance of Office Buildings Maintenance Management

Based on Figure 2, there are three maintenance elements that record medium correlation that is Tangibles, Internal Image and External Image with respective coefficients of 0.67, 0.61 and 0.62. Assurance (r=0.41, p>0.05) and Cleaning (r=0.33, p>0.05) signify low correlation between their levels of performance and systems. Meanwhile the other fourteen (14) maintenance elements that are also the variables record very low correlations for the respective correlations are between 0.01 and 0.30. Overall, it can be concluded that there is a minimal impact of correlation between the maintenance management performance and systems applied.

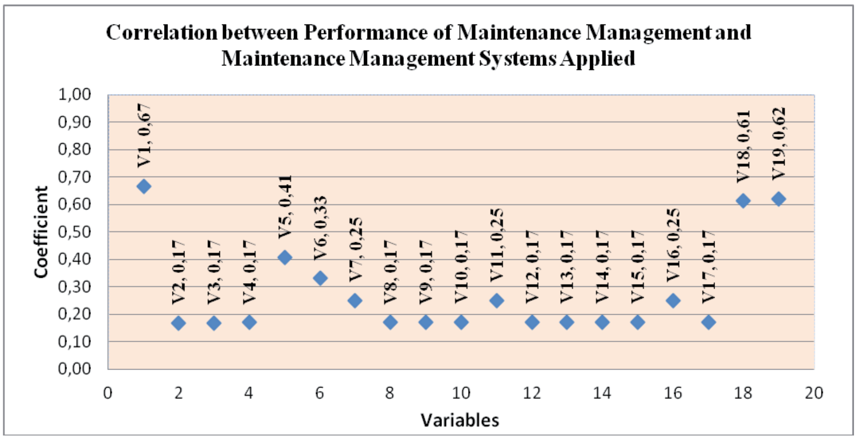


Fig.2. Performance of building image

7. Conclusion

This study has given an overview of the scenario of maintenance management of high-rise office buildings particularly on the development of maintenance management system and also performance measurement systems. Investigations on the maintenance management system and performance of maintenance management along with the relationship between them are accomplished. The findings suggest that in general the common maintenance management systems applied for office building comprises of three (3) major aspects namely Functional, Technical and Image. Important service elements for instance Service Characteristics, Building Services and Building Image encompass the three (3) major aspects respectively. This study has also found that generally all five (5) office buildings chosen have an average maintenance management performance as rated by respective end users. Findings also signify that in general background of the respondents has a significant relationship with the performance of maintenance management system. The most significant finding from the study is that there is a positive relationship between the maintenance management systems and performance of maintenance management especially in several elements of Service Characteristics and Building Services. Besides, it is also noted that maintenance managers have a similar perception in the importance ranking of maintenance management service elements with a fractionally difference of ranking order.

This survey has shown that benchmarking or assessment on the performance of maintenance management is very important as it enables the maintenance managers to comprehend the strengths, weaknesses and also significance of the service provided and also both tangible and intangible values of the building. Indirectly, maintenance managers can identify any probable threats or risks of their services. Concurrently, the establishment of maintenance management performance level is beneficial for the maintenance managers to implement immediate actions to improve the performance. It also serves as a signal that a major transformation is highly required to enhance the quality of performance. The positive relationship also ascertains that the implementation standard of maintenance management systems determine the performance of maintenance management system. At the same time, the difference shown in the priorities of maintenance management service elements signifies a strong emphasis on users' needs and requirements is required from maintenance managers.

In general, therefore, it seems that the study has achieved the aims of the study through a valid performance measurement design. The high value of Alpha's Cronbach validates that the variables used for the survey are reliable and the performance measurement design can be considered as a robust instrument to measure maintenance management performance.

8. Recommendations

It is strongly advised that maintenance managers should value the important roles of end users in evaluating the performance of maintenance services with a great attention given on their needs and requirements. Consultations with the end users should be a mechanism to establish a proactive management process. Maintenance managers must also consider implementing a continuous benchmarking or assessments on the services provided and subsequently focus on any critical service elements identified. A thorough analysis on the implementation of all maintenance services and respective sub-contractors helps to identify the weaknesses and criteria that need to be improved.

Besides, it is highly recommended that a maintenance management guideline is provided to standardize the maintenance standard practiced office building maintenance managers. In relation with this, a statutory act on the compliance of maintenance management system criteria and regulations should be established to improve the maintenance management performance and also to avoid any mismanagement which could result to corruptions and abuse of power.

Future research on the maintenance scope is most encouraged specifically on performance level of maintenance management on a larger scale, the implications of maintenance management failure, cost analysis of maintenance management, performance measurement assessment on all classes of residential housing and public buildings and a proposal on maintenance management statutory acts.

***Acknowledgement**:*

The authors gratefully acknowledge the financial support of the University of Malaya Research Grant (UMRG), grant no RG171/12SUS established at the University of Malaya, Sustainability Science Research Cluster.

References

1. Ahmad R. Role and Approach of Building Surveying in Facilities & Maintenance Management in National Seminar on Facilities & Maintenance Management In the 21st Century: The Demands & Needs for Growth In Malaysian Building Industry, Kuala Lumpur: 12th August 2008.
2. Al-Sultan KS, Duffuaa SO. Maintenance control via mathematical programming. Journal of Quality in Maintenance Engineering 1995; 1(3): 36–46.
3. Azizul A. The Importance of FMM towards Efficient Building Services Performance. National Seminar on Facilities & Maintenance Management in the 21st Century: The Demands & Needs for Growth In Malaysian Building Industry, Kuala Lumpur: 12th August 2008.
4. Bergman MM. On Concepts and Paradigms in Mixed Methods Research. Journal of Mixed Methods Research 2010; 4(3): 171–175.
5. Bernand TL, Payant RP. Facility Inspection Field Manual: A complete condition assessment guide. USA: McGraw-Hill Companies, 2000.
6. Certo S. Modern Management. New Jersey, USA: Prentice-Hall Inc., 2000.
7. Coetzee JL. Maintenance. Republic of South Africa: Maintenance Publishers, 1998.
8. Danielson B. A Study of Maintenance Problems in Swedish Mines, Study Report. Swedish: Idhammar Konsult AB, 1987.
9. Egbu C. Sustainability in Building Facilities & Maintenance: Skills, Knowledge and Personal Qualities for Managing Refurbishment Project. National Seminar on Facilities & Maintenance Management in the 21st Century: The Demands & Needs for Growth In Malaysian Building Industry, Kuala Lumpur, 2008.
10. Gelders L, Mannaerts P, Maes J. Manufacturing strategy, performance indicators and improvement programmes. International Journal of Production Research 1994; 32(4): 797-805.
11. Hinks J. Business Related Performance Measures for Facilities Management in K. Alexander (ed) Facilities Management Innovation and Performance. Oxon: Spoon Press; 97-109, 2004.
12. Johnson B, Onwuegbuzie A. Mixed method research: A research paradigm whose time has come. Educational Researcher 2006; 33(7): 14-26.
13. Kaplan RS, Norton DP. The balances scorecard: Measures that drive performance. Harvard Business Review 1992; January-February: 71-79.
14. Kutucuoglu KY, Hamali J, Irani Z, Sharp JM. A framework for managing maintenance using performance measurement systems. International Journal of Operations & Production Management 2001; 21(1): 173-195.
15. Leech NL, Onwuegbuzie AJ. A Typology of Mixed Methods Research Designs. Quality and Quantitative 2009; 43(2): 265–275.
16. Mertens DM. Mixed Methods as Tools for Social Change. Journal of Mixed Methods Research 2011; 5(3): 195–197.
17. Miles MB, Huberman AM. Qualitative Data Analysis (2nd edition). Sage: Thousand Oaks, CA, 1994.
18. Morgan DL. Paradigms Lost and Pragmatism Regained. Journal of Mixed Method Research 2007; 1(1), 48–76.
19. Myeda NE, Kamaruzzaman SN, Pitt M. Measuring the Performance of Office Buildings Maintenance Management in Malaysia. Journal of Facilities Management 2011; 9 (3):181–199.
20. Natasha K, Nawawi AH, Hashim AE, Husin HN. Performance Analysis of Government and Public Buildings via Post Occupancy Evaluation. Research in the Malaysian Construction Industry and Environment: Past, Present and Future held in Gombak. Conducted by the System Research Group. Gombak: Kulliyah of Architecture & Environmental Design International Islamic University: 2008.
21. Parida A, Kumar U. Maintenance Performance Measurement (MPM): Issues and Challenges. Journal of Quality in Maintenance 2006; 12(3): 239–251.
22. Schmberger R, Knod J. Operations Management: Customer Focus Principles, USA: Irwin, 1997.
23. Spires C. Asset and Maintenance Management-becoming a boardroom issue. Managing Service Quality 1996; 6(3); 13–15.
24. Tsang AHC, Hui EYY. Sourcing strategies of facilities management. Journal of Quality in Maintenance Engineering 2004; 10(2): 85–92.
25. Wordsworth P. Lee's building maintenance management: 4th Edition. Oxford: Blackwell Science Ltd., 2001.
26. Yasmin M. Recognising the Importance of Service Operations Aspects In Managing High-Rise Commercial Buildings In Kuala Lumpur: The Management Operations Concepts. The Malaysian Surveyor 2003; 38(3): 37–42.

Associate Professor Dr. Sr Syahrul-Nizam KAMARUZZAMAN
Department of Building Surveying,
Faculty of Built Environment,
University of Malaya, 50603 Kuala Lumpur, Malaysia
E-mail: syahrulnizam@um.edu.my

Nik Elyna MYEDA
Bartlett School of Graduate Studies
University College London
Central House, 14 Upper Woburn Place
London WC1H 0NN, United Kingdom
E-mail: elyna.nik.10@ucl.ac.uk

Professor Dr. Michael PITT
Bartlett School of Graduate Studies
University College London
Central House, 14 Upper Woburn Place
London WC1H 0NN, United Kingdom
E-mail: michael.pitt@ucl.ac.uk

Article citation info:
LI Y-F, MI J, HUANG H-Z, XIAO N-C, ZHU S-P. System reliability modeling and assessment for solar array drive assembly based on bayesian networks. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 117–122.

Yan-Feng LI
Jinhua MI
Hong-Zhong HUANG
Ning-Cong XIAO
Shun-Peng ZHU

SYSTEM RELIABILITY MODELING AND ASSESSMENT FOR SOLAR ARRAY DRIVE ASSEMBLY BASED ON BAYESIAN NETWORKS

MODELOWANIE I OCENA NIEZAWODNOŚCI SYSTEMU W OPARCIU O SIECI BAYESOWSKIE NA PRZYKŁADZIE UKŁADU NAPĘDU PANELI SŁONECZNYCH

Along with the increase of complexity in engineering systems, there exist many dynamic characteristics within the system failure process, such as sequence dependency, functional dependency and spares. Markov-based dynamic fault trees can figure out the modeling of systems with these characteristics. However, when confronted with the issue of state space explosion resulted from the growth of system complexity, the Markov-based approach is no longer efficient. In this paper, we combine the Bayesian networks with the dynamic fault trees to model the reliability of such types of systems. The inference technique of Bayesian network is utilized for reliability assessment and fault probability estimation. The solar array drive assembly is used to demonstrate the effectiveness of this method.

Keywords*: fault tree, dynamic fault tree, Bayesian network, system reliability, solar array drive assembly.*

Wraz ze wzrostem złożoności w systemach technicznych, pojawia się wiele charakterystyk dynamicznych w ramach procesu awarii systemu, takich jak zależność sekwencyjna, zależność funkcjonalna czy zabezpieczające elementy zapasowe. Oparte na koncepcjach Markowa dynamiczne drzewa uszkodzeń mogą posłużyć do modelowania systemów z powyższymi charakterystykami. Jednak w konfrontacji z problemem eksplozji stanów wynikającym ze wzrostu złożoności systemu, podejście oparte na teoriach Markowa nie jest już skuteczne. W niniejszej pracy łączymy sieci bayesowskie z dynamicznymi drzewami uszkodzeń w celu modelowania niezawodności tego typu systemów. Technikę wnioskowania sieci bayesowskiej wykorzystano do oceny niezawodności i prawdopodobieństwa wystąpienia uszkodzenia. Skuteczność niniejszej metody wykazano na przykładzie układu napędu paneli słonecznych.

Słowa kluczowe*: drzewo uszkodzeń, dynamiczne drzewo uszkodzeń, sieć bayesowska, niezawodność systemu, układ napędu paneli słonecznych.*

1. Introduction

Along with the increasing complexity of the structure and function of modern military satellites, meteorological satellites, commercial broadcasting and telecommunication satellites, there have been many new requirements for their reliability that have been put forward. The system is anticipated to achieve the required function and the high reliability while operating in a harsh environment. Modern satellites generally utilize solar power, a kind of sustainable energy, to meet the energy requirements and to sustain a long service lifetime. A array of solar batteries is installed in the satellite to produce electric energy. To generate enough energy, a drive assembly is utilized to rotate the solar array that directly faces the solar beam to receive the maximum solar irradiance. Therefore, a comprehensive and rigorous reliability analysis for solar array drive assembly is of great importance.

The traditional fault tree analysis method is a system reliability analysis methodology that is based on two states and static fault logic. It has been widely used for reliability and safety analysis in complex engineering systems. However, for many systems, there exist many complex dynamic characteristics during the failure process of modern engineering systems, such as sequence dependency and functional dependency. Static fault trees are not capable of modeling these failure processes.

Dugan presented a dynamic fault tree (DFT) method for analyzing reliability of systems with these aforementioned dynamic characteristics [2]. In this method, sets of dynamic gates are defined first for modeling these dynamic failure mechanisms. Then, the DFT model is decomposed into independent static modules and dynamic modules by functional decomposition, where Binary Decision Diagram (BDD) is employed to solve the static module and Markov Chain resolves the dynamic module issue.

As system dimensions increase, the number of components and the failure logic between them are become extremely large and complex. As the result, the determination and quantification of an appropriate Markov model increase exponentially. Thus, the efficiency of a Markov model has led to various concerns by many reliability analysts and experts. Amari presented a numerical integration based method for calculating the probability of dynamic gates [8]. It can alleviate the state space explosion issue encountered by Markov models. However, it fails to calculate the importance of basic components as well as the reverse inference. Rao proposed a Monte Carlo simulation based DFT method [1]. Yuge presented an inclusion-exclusion principle based method for solving DFT [3]. It can precisely calculate the probability of occurrence of the top event of fault trees with Priority-AND gate and repeated basic events given the minimal cut sets. However, they also cannot calculate the importance measure of basic events.

Bayesian networks (BN) captures the nodal relationships using a graphic approach [9]. By utilizing the conditional independency between nodes, it reduces the dimensions of the conditional probability table of non-root nodes as well as the complexity of all corresponding calculations. By adding different kinds of evidence into the BN, we can perform the forward reliability assessment, conduct the backward fault diagnosis, and further calculate the importance of individual components. Boudali proposed a discrete time BN based methodology for system reliability modeling and assessment [4].

The remainder of the paper is organized as follows. Section 2 provides a brief introduction on BN and its inference capability. In Section 3, we present the discrete time BN model, and outline the quantification of conditional probability table of static and dynamic logic gates. Section 4 presents an application example. We summarize our work and conclusion in Section 5.

2. BN model

2.1. Review of BN and conditional independency

A BN is a directed acyclic graph comprised of nodes and arcs. Nodes represent random variables (RVs) and the arcs between pairs of nodes capture the dependency information between the RVs [5]. Each root node has a prior probability table (PPT) representing the probability distribution of the nodes in its state space. A non-root node has a conditional probability table (CPT) representing its conditional probability distribution under the state combination of its father nodes [5,7].

Consider a system with 5 components, denoted as A, B, C, D, and E. We further assume that components and the system have two states. Without considering the conditional independency, the expression of joint probability distribution of these 5 variables can be given as:

$$P(A,B,C,D,E) = P(A)P(B \mid A)P(C \mid A,B)P(D \mid A,B,C)P(E \mid A,B,C,D) \tag{1}$$

The number of independent parameters is $31(2^5-1)$.

Suppose that A is independent of B, D is independent of A and B given C, and E is independent of A, B, and D conditional on C. The decomposition of the joint probability distribution of these variables is:

$$P(A,B,C,D,E) = P(A)P(B)P(C \mid A,B)P(D \mid C)P(E \mid C) \tag{2}$$

From Eq. (2), the number of independent parameters is reduced to 10, which significantly reduces the dimensions and complexity of calculation.

2.2. An example of BN and its bidirectional reasoning

In order to explain the principles of BN fully, an example of BN is shown in Fig. 1. It contains four root nodes, one intermediate node and one leaf node. As mentioned above, each node has a PPT or CPT to quantify its probability distribution given the state combination of its parent nodes. Utilizing the inference algorithm of BN, we can implement the probability inference between these variables. The probability distribution of root nodes is given in Fig. 1.

Utilizing the joint tree inference algorithm of BN, the probability distribution of node T without evidence can be calculated:

$$P(T=0)=0.9793,\ P(T=1)=0.0207$$

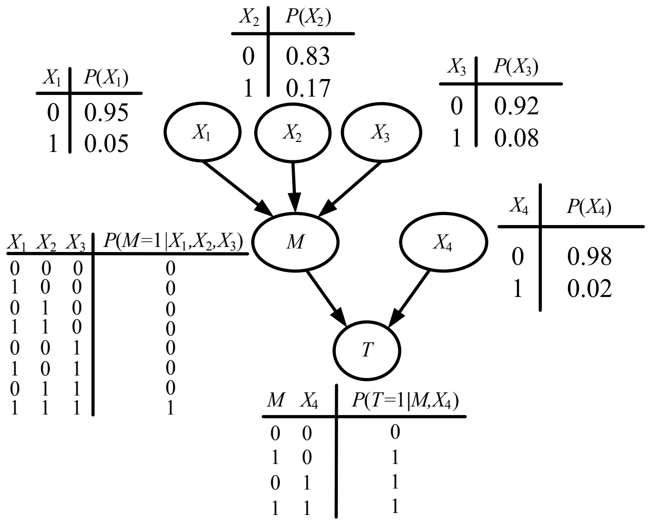


Fig. 1. An example of BN

Supposing each root node is in state 1 respectively, the conditional probability distribution of leaf node T is shown in Table 1.

Table 1. Conditional probability distribution of leaf node T

X_i	X_1	X_2	X_3	X_4
$P(T=0\|X_i=1)$	0.9667	0.9761	0.9717	0.0000
$P(T=1\|X_i=1)$	0.0333	0.0239	0.0283	1.0000

Supposing the leaf node is in state 1, the posterior probability distribution of root nodes is shown in Table 2.

Table 2. The posterior probability distribution of root nodes

X_i	X_1	X_2	X_3	X_4
$P(X_i=0\|T=1)$	0.9194	0.8032	0.8903	0.0322
$P(X_i=1\|T=1)$	0.0806	0.1968	0.1097	0.9678

By utilizing the bidirectional inference, we can implement probability inference between nodes within the BN and perform the reliability analysis and probability inference of each potential fault in engineering systems.

3. Reliability assessment model based on DFT and discrete time BN

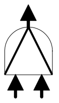
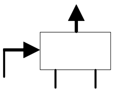
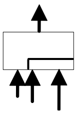
3.1. DFT

DFT methodology extends traditional fault tree method by defining a set of dynamic logic gates to represent the time dependency and functional dependency among component failure mechanisms. Priority-AND gate (PAND), Functional Dependent Gate (FDEP) and Spare are most commonly used dynamic gates. Their notation and failure mechanism are listed in Table 3.

3.2. Discrete time BN model for system reliability modeling and assessment

Consider the reliability of a system at mission time t. The time interval $(0,t]$ is evenly divided into n equal intervals, the length of each interval is $\Delta=T/n$. Let the interval $(t, +\infty)$ be the $(n+1)$th interval. So the time line is divided into $n+1$ intervals. We define the $n+1$ intervals

Table 3. Dynamic gates and failure mechanism

Dynamic gate	Notation	Failure mechanism
PAND		If and only if both A and B occur, and A occurs before B, the output occurs
FDEP		If the trigger event occur, all the dependent events occur
Spare		The primary input is initially powered on, and all spares are standby. While primary input fail, the defective items are replaced one by one. If all inputs fail, the output event occurs.

as the state space of nodes in BN, and marked as $1,2,\cdots,n+1$, respectively. Then, if a node is in state i, it means that it will fail in time interval $((i-1)\Delta, i\Delta]$. The failure time of components and the system corresponds to the state of node in BN. If a node or component is in state n+1, it means that it will not fail at mission time t. The probability of the system being in state n+1 represents the reliability of the system at time t. The CPT of non-root node will be discussed in the next section.

3.3. Determination of CPT in BN corresponding to various gates in DFT

Before the quantification of system reliability, the DFT model is needed to transform into a BN model. For a given BN, the graphic structure represents the qualitative relationship between an individual node and its parent and child nodes. Meanwhile the CPT with each non-root node represents the quantitative relationship between the same nodes. Generally, it is quite easy to acquire the graphic structure of the BN according to the structure of DFT. Next, we will discuss the CPD of various kinds of gates, which could be transferred to CPT in its corresponding BN models.

(1) AND Gate

Let $X=[X_1,X_2,\cdots,X_m]$, where $X_i, i=1,2,\cdots,m$ denote the state variable of input events and m is the number of input events to the gate. Let Y be the state variable of the output event. The state space of each variable is $\{1,2,\cdots,n+1\}$. Let $k=\max(X_1,X_2,\cdots,X_m)$. The conditional probability distribution of Y under a certain state combination of input events can be given by

$$P(Y=j\mid X)=\begin{cases}1, & j=k\\ 0, & j\neq k\end{cases} \tag{3}$$

(2) OR Gate

The notation and meanings of all variables and their corresponding state spaces are the same as aforementioned AND Gate. Let $r=\min(X_1,X_2,\cdots,X_m)$. The conditional probability distribution of output of OR Gate is

$$P(Y=j\mid X)=\begin{cases}1, & j=r\\ 0, & j\neq r\end{cases} \tag{4}$$

(3) Priority-AND Gate

Let A and B be the input events of the Priority-AND gate, Y is the output event, and a, b, y are their values, respectively. The conditional probability distribution of Y is

$$\text{While } a<b\le n+1,\ P(Y=i)=\begin{cases}1, & i=b\\ 0, & \text{else}\end{cases} \tag{5}$$

$$\text{While } a\ge b,\ P(Y=i)=\begin{cases}1, & i=n+1\\ 0, & \text{else}\end{cases} \tag{6}$$

(4) Functional Dependent Gate

For the case that the functional dependent gate has only one trigger event A, the CPT of the output B dependent of A on condition that A has happened will be an identity matrix. Its CPD is

$$P(B=j\mid A=i)=\begin{cases}1 & j=i\\ 0 & j\neq i\end{cases} \qquad i,j=1,2,\cdots,n+1 \tag{7}$$

Meanwhile the FDEP has two or more trigger events, and an intermediate node is inserted between the output node and all the trigger nodes. The CPT of the intermediate node depends on the relationship between the intermediate node and all the trigger nodes. The CPT of the output node will also be an identity matrix as it is only dependent on the intermediate node.

(5) Cold Spares

According to the failure mechanism of Cold Spare (CSP), the failure distribution of the spare is related to its failure rate and the failure time of the primary component. Let x,y be the state of primary input A and spare B, and λ is the failure rate of A and B. Thus, the conditional probability distribution of B is

$$P(B=y\mid A=x)=\begin{cases}\dfrac{\int_{(x-1)\Delta}^{x\Delta}\int_{(y-1)\Delta}^{y\Delta}\lambda e^{-\lambda(t-\tau)}\lambda e^{-\lambda\tau}\,d\tau\,dt}{\int_{(x-1)\Delta}^{x\Delta}\lambda e^{-\lambda\tau}\,d\tau}=\lambda\Delta e^{\lambda x\Delta}e^{-\lambda y\Delta} & ,x<y<n+1\\ 1-\sum\limits_{i=x+1}^{n}P(B=i\mid A=x) & ,x<y=n+1\\ 0 & ,x\ge y\end{cases} \tag{8}$$

(6) Hot Spares

In spite of the difference in failure mechanisms between AND Gate and Hot Spare Gate (HSP), the conditional probability distribution of the output nodes for hot spares are identical, so they are not repeated again.

4. Reliability modeling and assessment of solar array drive assembly of satellite based on BN

The solar array drive assembly (SADA) plays a vital role in the orientation function of a satellite solar array. Therefore, it is significant to perform reliability analysis and assessment on the drive assembly to ensure that the high reliability performance of the solar array and the entire satellite system. This paper presents a reliability analysis and assessment method based on the DFT and BN methodology.

The SADA provides the linkage between the solar array and the satellite body, with the function of driving the array to rotate to parallel with the solar beam to get as more solar energy as possible. The studied SADA is comprised of solar array sensitive apparatus, on-

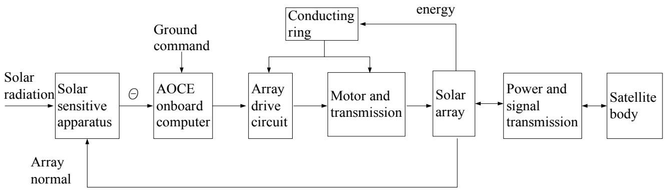


Fig. 2. Working principle diagram of orientation system of SADA

board computer, conducting ring, drive circuit, and motor transmission and so on. Its working principle diagram is shown in Fig. 2 [6].

In the paper, the event 'Failure of SADA' is defined as the top event. Let A, B, C, D, E, F, G, H, I, S, and K represent failure events of solar sensitivity apparatus, on-board computer, harmonic reducer, drive motor, human factor, electrical system, transmission, conducting ring, position sensor, windings and electric brush, respectively. Let A_1 and A_2, B_1 and B_2, S_1 and S_2, F_1 and F_2, I_1 and I_2, K_1 and K_2 be the primary and spare of solar sensitivity apparatus, on-board computer, windings, electrical system, position sensor and brush, respectively.

The solar sensitivity apparatus, on-board computer, stator windings of motor and position sensor can be modeled by the CSP gate. The electrical system can be modeled using FDEP gate as it has components with functional dependency characteristics. The conducting ring is modeled using HSP gate.

Suppose that all other failures between components are independent apart from aforementioned dependent failures. Finally, we get the DFT of SADA shown in Fig. 3. After mapping all the events of the DFT into the nodes of BN, the BN model of SADA could be obtained as shown in Fig. 4.

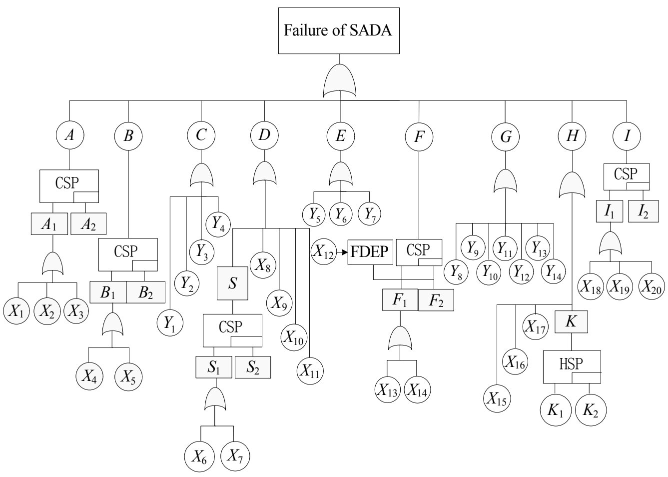


Fig. 3. DFT of SADA

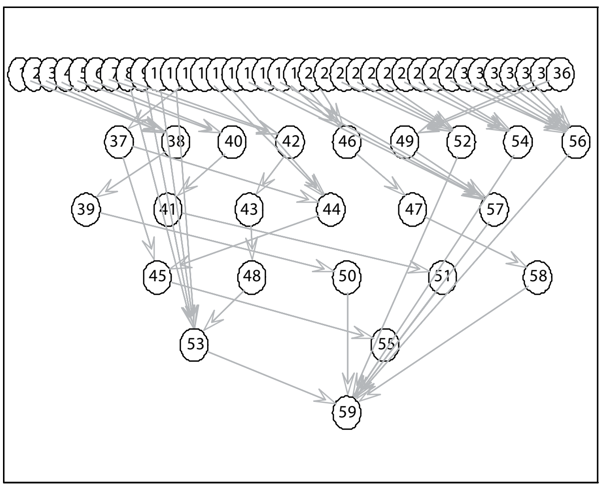


Fig. 4. BN model of SADA

The code, name, and failure rate of all basic events are listed in Table 4. The system reliability curve for mission time 50000 hours when n was taken 2, 3, and 4 is shown in Fig. 5. The system reliability for every 100 hours when n were taken 2, 3, 4 under [45000, 50000] hour mission time is shown in Fig. 6. From the statistical analysis based on the calculated data, we know that the maximum error of reliability between the second set data and the first set data is 0.0606%, meanwhile the maximum error between the third set data and the second set data is 0.0305%. Taking into account the existence of the other uncertainties in the system, n can be taken to be 4 in order to meet the error requirements for reliability.

Table 4. Code, name and failure rate of basic events (10-6failure/hour)

Code	Event name	Failure rate (λ)	Code	Event name	Failure rate (λ)
X_1	Failure of optical head	0.5	X_{19}	Failure of LEDs	0.5
X_2	Censer failure	0.6	X_{20}	Failure of detection circuit	0.5
X_3	Failure of signal processing circuit	0.12	Y_1	Failure of flexible gear	0.1
X_4	Hardware failure of on-board computer	0.5	Y_2	Gear wear	0.6
X_5	Software failure of on-board computer	0.25	Y_3	Failure of solid lubricant film	0.55
X_6	Fatigue failure of windings	0.25	Y_4	Failure of lubricating grease	0.5
X_7	Windings are burned down	0.1	Y_5	Go fly of instruction	0.5
X_8	Failure of drive circuit	0.12	Y_6	Mistakes in hardware design	0.2
X_9	Locked rotor	0.12	Y_7	Mistakes in software design	0.2
X_{10}	Friction increases	0.1	Y_8	Clutch failure	0.6
X_{11}	Fatigue failure	0.1	Y_9	The bearing is jammed	0.5
X_{12}	Interface fault	0.6	Y_{10}	Bond breakage	0.1
X_{13}	Circuit failure	0.125	Y_{11}	Failure of Rotation axis of potentiometer	0.5
X_{14}	Transistor failure	0.5	Y_{12}	Gear failure	0.1
X_{15}	The circuit is burned down by discharge	0.1	Y_{13}	External failure	0.5
X_{16}	Failure of bearing lubricant	0.1	Y_{14}	Comprehensive failure	0.5
X_{17}	Failure of insulation	0.15	K_1	Primary failure of electric brush	0.5
X_{18}	Failure of image censer	0.5	K_2	Spare failure of electric brush	0.5

The probability distributions of the top event T for 5 time interval under t=50000, n=4 are shown in Table 5. That is, the system probability of failure and reliability within the mission time are 0.2931 and 0.7069, respectively.

When the system in the failure state (the state of leaf node T is 5), the probabilities of failure for every bottom event can be calculated by using the reverse inference of the BN as shown in Table 6. From

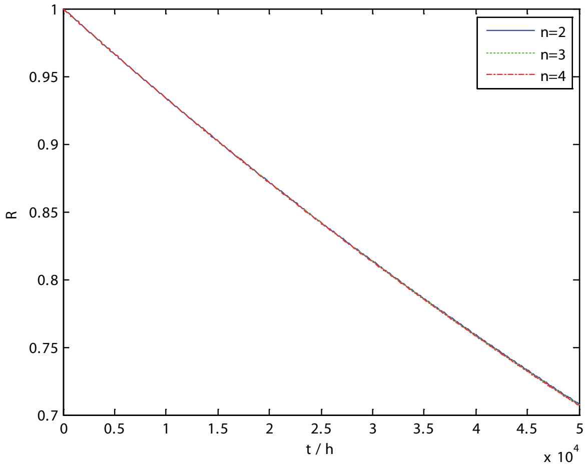


Fig. 5. System reliability curve for mission time 50000 hours

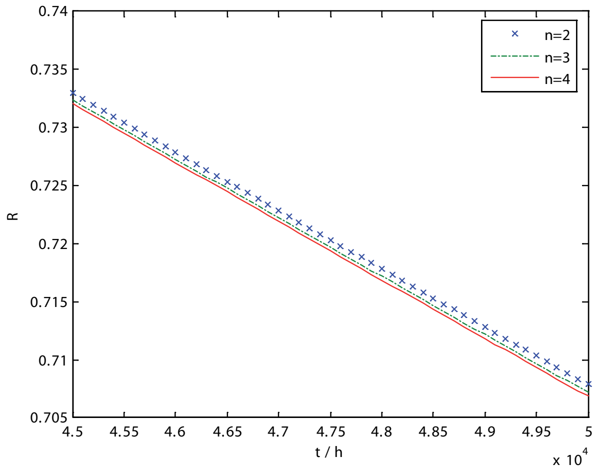


Fig. 6. Reliability comparisons for n at a given value 2, 3 and 4, respectively

Table 5. Probability distributions of top event under n=4

$T=i$	1	2	3	4	5
$P(T=i)$	0.0820	0.0759	0.0703	0.0649	0.7069

Table 6 we can conclude that the minimum probability of failure is X_7. The maximum probability of failure is X_{28} , which is also the weakest part of the system.

Based on the reasoning algorithm of the BN, the failure conditional probabilities of the top event on condition that the bottom event is on failure state can be calculated as shown in Table 7.

Table 6. Probability of failure for every component under system failure

X_i	$P(X_i=1\|T=1)$	X_i	$P(X_i=1\|T=1)$	X_i	$P(X_i=1\|T=1)$
X_1	0.0260	X_{13}	0.0064	Y_5	0.0842
X_2	0.0311	X_{14}	0.0254	Y_6	0.0339
X_3	0.0063	X_{15}	0.0170	Y_7	0.0339
X_4	0.0255	X_{16}	0.0170	Y_8	0.1008
X_5	0.0128	X_{17}	0.0255	Y_9	0.0842
X_6	0.0126	X_{18}	0.0262	Y_{10}	0.0170
X_7	0.0051	X_{19}	0.0262	Y_{11}	0.0842
X_8	0.0204	X_{20}	0.0262	Y_{12}	0.0170
X_9	0.0204	Y_1	0.0170	Y_{13}	0.0842
X_{10}	0.0170	Y_2	0.1008	Y_{14}	0.0842
X_{11}	0.0170	Y_3	0.0925	K_1	0.0261
X_{12}	0.1008	Y_4	0.0842	K_2	0.0261

Table 7. System probabilities of failure on condition that each bottom event is in failed state

X_i	$P(T=1\|X_i=1)$	X_i	$P(T=1\|X_i=1)$	X_i	$P(T=1\|X_i=1)$
X_1	0.3083	X_{13}	0.3012	Y_5	1.0000
X_2	0.3083	X_{14}	0.3011	Y_6	1.0000
X_3	0.3084	X_{15}	1.0000	Y_7	1.0000
X_4	0.3027	X_{16}	1.0000	Y_8	1.0000
X_5	0.3027	X_{17}	1.0000	Y_9	1.0000
X_6	0.2977	X_{18}	0.3115	Y_{10}	1.0000
X_7	0.2977	X_{19}	0.3115	Y_{11}	1.0000
X_8	1.0000	X_{20}	0.3115	Y_{12}	1.0000
X_9	1.0000	Y_1	1.0000	Y_{13}	1.0000
X_{10}	1.0000	Y_2	1.0000	Y_{14}	1.0000
X_{11}	1.0000	Y_3	1.0000	K_1	0.3102
X_{12}	1.0000	Y_4	1.0000	K_2	0.3102

5. Conclusions

This study proposed a hybrid reliability modeling and assessment method based on Bayesian network and dynamic fault tree. The method for determining the conditional probability of logic gates is investigated as well. The dynamic fault tree model and its corresponding Bayesian network are tested and applied to predict the lifetime of solar array drive assemblies. The junction tree inference algorithm is used for the proposed hybrid model. The results can be used for system failure diagnosis and is expected to identify and further remove the design weakness of the system. The results obtained from the satellite's solar power control system have shown that the proposed method is effective and accurate, showing its potential application for reliability assessment on large and complex engineering systems.

Acknowledgements

This research was partially supported by the National Natural Science Foundation of China under the contract number 51075061 and the National Programs for High Technology Research and Development of China under the contract number 2007AA04Z403.

6. References

1. Amari S, Dill G, Howald E. A new approach to solve dynamic fault trees. In: Annual IEEE Reliability and Maintainability Symposium, 2003. p. 374-379.
2. Boudali H, Dugan JB. A discrete-time Bayesian network reliability modeling and analysis framework. Reliab Eng Syst Saf 2005; 87(3): 337–349.
3. Dugan JB, Bavuso SJ, Boyd MA. Dynamic fault-tree for fault-tolerant computer systems. IEEE Trans Reliab 1992; 41(3): 363–76.
4. Rao KD, Gopika V, Rao VVSS, et al. Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. Reliab Eng Syst Saf 2009; 94(4): 872–883.
5. Langseth H, Portinale L. Bayesian networks in reliability. Reliab Eng Syst Saf 2007; 92(1): 92–108.
6. Shi L. Reliability analysis of solar array drive assembly. Chengdu: University of Electronic Science and Technology of China, 2011.
7. Wilson AG, Huzurbazar AV. Bayesian networks for multilevel system reliability. Reliab Eng Syst Saf 2007; 92(10): 1413–1420.
8. Yuge T, Yanagi S. Quantitative analysis of a fault tree with priority AND gates. Reliab Eng Syst Saf 2008; 93(11): 1577–1583.
9. Zhang LW, Guo HP. Introduction to Bayesian networks. Beijing: Science Press, 2006.

Yan-Feng LI, Ph.D. candidate
Jinhua MI, Ph.D. candidate
Prof. Hong-Zhong HUANG, Ph.D.
Ningcong XIAO, Ph.D. candidate
Shun-Peng ZHU, Ph.D.
School of Mechanical, Electronic, and Industrial Engineering
University of Electronic Science and Technology of China
No. 2006, Xiyuan Avenue, West Hi-Tech Zone, Chengdu, Sichuan, 611731
P. R. China
E-mail: hzhuang@uestc.edu.cn

Article citation info:
TARASIUK P, MOLSKI KL, SZYMANIUK A. Fatigue designing of welded agricultural wheels. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 123–128.

Piotr TARASIUK
Krzysztof Lech MOLSKI
Andrzej SZYMANIUK

FATIGUE DESIGNING OF WELDED AGRICULTURAL WHEELS

PROJEKTOWANIE TRWAŁOŚCIOWE SPAWANYCH KÓŁ POJAZDÓW ROLNICZYCH*

Each introduction of new or modified products into production requires conducting a series of calculations and experiments which would verify the desired quality. In case of agricultural wheels it is necessary to estimate their fatigue life and durability at a possibly low cost. This paper presents a method of fatigue design of agricultural wheels on the example of a welded wheel type 9.00x15.3. A previously designed numerical model FEM was used, which enables to identify the potential weak spots of the construction which could determine its total durability and allows to assess the critical parameters. Based on the results of model studies, construction changes were introduced, which were experimentally verified on a fatigue machine during durability testing for radial forces.

Keywords*: disc wheels, fatigue design, Hot-Spot.*

Wprowadzenie do produkcji nowych lub zmienionych pod względem konstrukcyjnym wyrobów wymaga wykonania obliczeń sprawdzających oraz doświadczeń potwierdzających uzyskanie ich założonej jakości. W przypadku kół pojazdów rolniczych konieczne jest wcześniejsze oszacowanie ich wytrzymałości i trwałości zmęczeniowej możliwie niskim kosztem. W pracy przestawiono metodę projektowania trwałościowego kół pojazdów wolnobieżnych na przykładzie spawanego koła typu 9.00x15.3. Wykorzystano opracowany wcześniej model numeryczny MES, który umożliwia identyfikację potencjalnie najsłabszych miejsc konstrukcji decydujących o jej trwałości oraz pozwala określić wartości parametrów krytycznych. Na podstawie wyników badań modelowych wprowadzono zmiany konstrukcyjne, które zweryfikowano doświadczalnie na stanowisku badawczym w testach trwałościowych na obciążenia promieniowe.

Słowa kluczowe*: koła tarczowe, projektowanie trwałościowe, Hot-Spot.*

1. Introduction

The creation of new constructions or modification of existing ones must stem from an economic analysis and market requirements. In the case of agricultural wheels the necessity of increasing the quality of construction is caused by a relatively high share of material costs in the final product and a relatively low durability with regards to expectations. During usage, wheels undergo various types of strain, which may result in damage to the wheel. Due to constant operation, typical usage damages are a result of material fatigue and appear in some characteristic spots referred to as *theoretical critical zones TCZ* [12, 17]. During usage unique situations may occur, such as wheel collisions with an obstacle or a sudden puncture or tearing of the tyre, which may cause a permanent deformation of the wheel. Such cases will not be taken into account in this paper.

The fatigue life of every individual construction should be initially determined at the designing stage. The basic task of a constructor is then to identify the location of *TCZ*s and to assess their fatigue life for assumed usage conditions and with known material parameters. The appropriate norms, such as [4, 13] require the wheel durability to be verified in experiments with the use of fatigue machines. Obtaining the expected improvement in quality with a trial and error method is time-consuming, and prolonged analysis of wheels may be costly. It is therefore crucial to introduce construction analysis on models with use of computer technology, conducting fatigue tests in the last stages of work.

In this context *fatigue design of wheels* means a process of conscious and purposeful creation of a product characterised by the required fatigue life, based on the assumptions made, with use of theoretical knowledge in the field of material technology, methods of numerical modelling, fatigue life, fracture mechanics, as well as practical knowledge, while taking into account the influence of the technological process on durability.

The aim of this paper is to present the method of *fatigue design* used in welded agricultural wheels, aiming at obtaining more beneficial fatigue life with no increase in mass and with maintaining the same production technology. The method is presented on the example of a welded wheel type 9.00x15.3 with an *Implement* brand tyre [8], commonly used in agricultural vehicles.

2. Numerical modelling of the wheel and its experimental verification

The numerical model of agricultural wheels was developed earlier and described in [19]. The tyre played a key role in this model [9], which was described with the use of the hyperelastic Mooney-Rivlin material [11, 15], taking into account the five construction zones of different material properties. An example of the modelled wheel's geometry with an *Implement* type tyre and the division into individual zones is presented in figure 1.

In the correct wheel description, the various zones of the tyre presented in Fig. 1 need to be included, such as: the tread 1a, sidewall 1b, bead 1c, layer elements 1d and the steel-rubber bead bundle 2. The picture also presents the elements of the metal part of the wheel: rim 3, disc 5 and the weld 4. The values of the material parameters characterising the individual zones of the tyre were defined in accord-

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

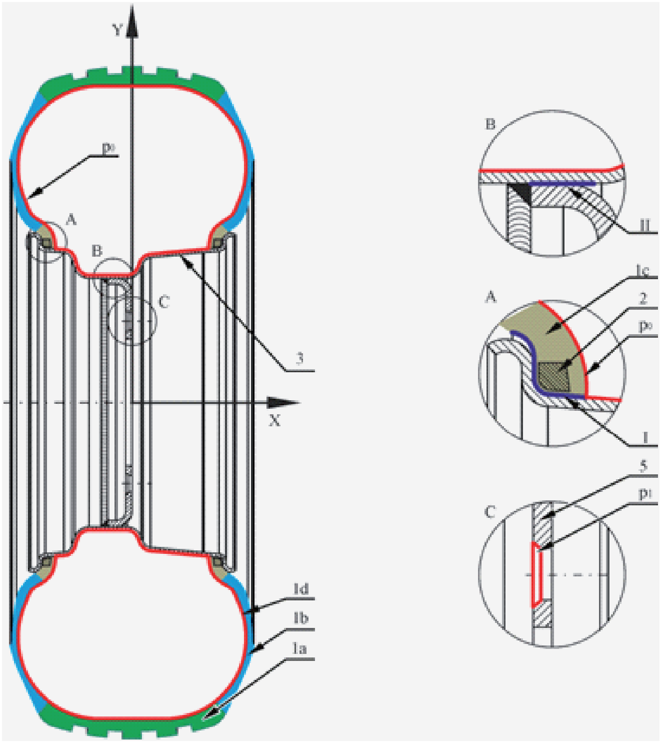


Fig. 1. Graphic scheme of the wheel model divided into the basic zones with characteristic construction properties. Description in the text

ance with the source data found in [5], and in case of the steel-rubber element (No. 2 in Fig. 1) the replacement parameters were determined as the harmonic mean of the material constants of the bead bundle and rubber, taking into account their percentage amount in the cross section.

The accuracy of the numerical calculations, proving the model's correctness, was verified in experiments with the use of the strain gauge method. Fifteen strain gauges were placed on the inner and outer part of the wheel's rim as well as on the outer part of the disc. After fitting the tyre and resetting the measuring device, the wheel was filled with air, fastened with screws on the workstation and gradually strained radially until the desire d force was obtained. During this time, the measuring device indications were being registered continuously. The maximum differences in deformation value calculated numerically and those acquired in strain gauge tests did not exceed 5%.

In Fig. 2 we can see the rim of the wheel with the strain gauges in place before the tyre was fitted. They were located in crucial places from the construction durability point of view, after an initial analysis of the numerical calculations results obtained by *FEM* for the designed model of a wheel strained radially.

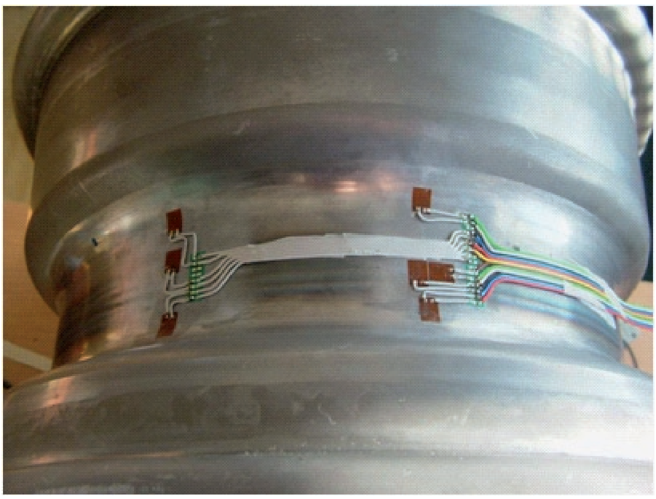


Fig. 2. The rim of the wheel with strain gauges in place from the tyre side

3. Identification of *TCZ* zones and their parameters

The results of fatigue tests of the wheels, conducted on a full scale at a fatigue machine (a running machine) and model numerical tests allowed to identify the actual location of construction damages and enabled to verify the used *FEM* models. Examples of fatigue damage and stress values acquired from the *FEM* solution are presented in Fig. 3. In the computer simulation, the wheel with a four-arm disc was pressed down to the surface with a force of P=26 kN in the radius direction, which corresponded to the conditions of fatigue tests conducted on the fatigue machine. The fatigue fracture visible in Fig. 3a) was located in the end zone of the weld, where a local increase in stress was noted, caused by the notch effect.

From the point of view of stress concentration and fatigue durability, the key role was that of the local geometrical qualities of the construction, such as: thickness, radius of curve and angles of the metal sheets, welded joints and notches located in the area of the weld cap and root.

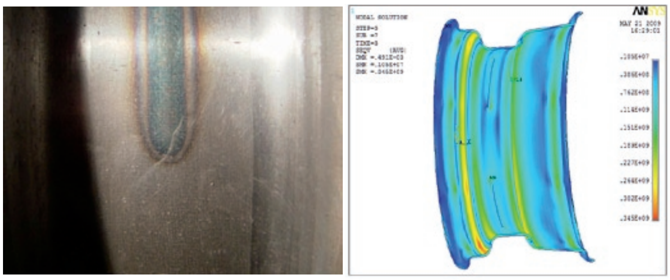


Fig. 3. a) Fatigue fracture at the end of the weld, b) distribution of stress in the FEM model of the rim with radial force applied

Basing on the numerical calculations, resulting from model analysis, three zones of increased concentration of stress were identified, which are a potential weak durability links (Fig. 4), namely: the curving of the wheel rim (*TCZ1*), the boundary of weld cap and rim material (*TCZ2*) and the area of the weld root joining the rim and disc (*TCZ3*).

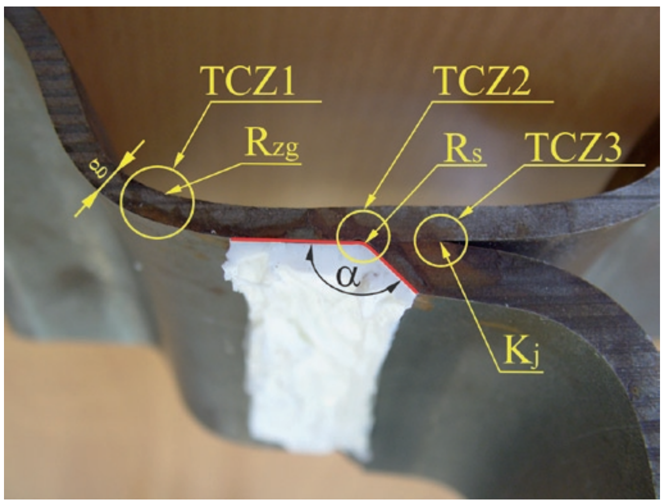


Fig. 4. A cross section of a weld joint with the characteristic values, areas of increased stress concentration marked and a layer of silicone enabling the identification of the local parameters of the notch

The characteristic local geometric features of the rim may be measured on real objects, whereas the parameters of weld joints may be conveniently identified for example using the method of silicone surface replicas.

Fig. 4 represents a cross section of a welded joint connecting the disc with the rim of the wheel. The characteristic values have been marked, such as the thickness of the sheet *g*, shape and radius of the curve R_{zg} and potential critical zones identified for the wheel model under radial load. The white silicone cast visible in the picture enabled the identification of the R_s value and the α angles. The results of these measurements are represented in figures 5 and 6.

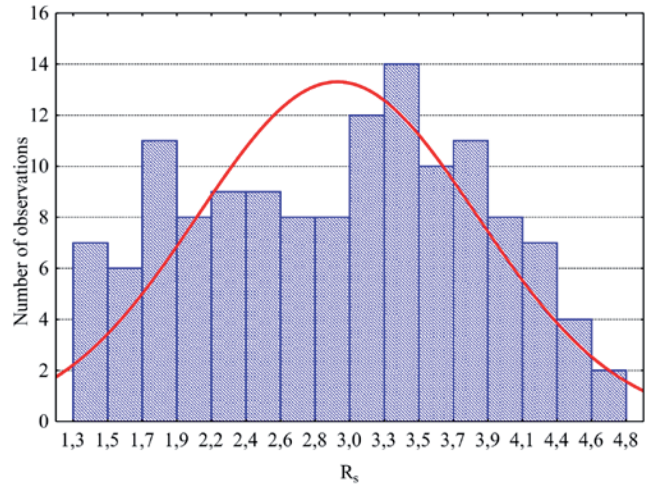


Fig. 5. Distribution of measured values R_s in a wheel with a four-arm disc

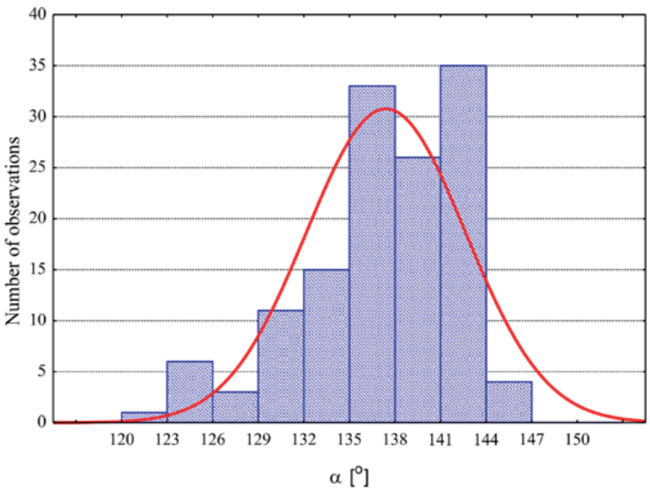


Fig. 6. Distribution of measured values of α angle in a wheel with a four-arm disc

In both presented cases the results of measurements may be described statistically by means of normal distribution. The average value of Rs radius was app. 3.0 mm with a standard deviation of 0.8 mm, whereas the average value of the α angle was 137° with a standard deviation of 6°.

The size of the R_{zg} radius and the thickness of the rim g are decisive in the value of local stress in the area *TCZ1* of the rim. By means of the Finite Elements Method the value and evolution of critical stress was determined in this area with a circumferential shift of the location of the reference point, starting from a location closest to the area of application of the radial force *P*. The character of these changes in points lying along the circumference is shown in Fig. 7. A characteristic feature of these stresses is their symmetry to the plane of force applied (coordinates '0' and '0.5' on the graph) and that the minimal stresses do not occur on the rim directly opposite to the area of application of the external force, but are located at an angle equal to appr. 1/3 of the wheel's revolution. This has a crucial meaning while calculating the number of cycles of stress changes in these areas, as well as the range and level of mean fatigue stresses in this area during a full revolution of the wheel around its axis.

In the location marked as *TCZ2* the value of stress needs to be calculated on the border of weld cap and rim material. In this case the Hot-Spot method is often used [6, 10, 14, 16], which is presented in Fig. 8.

This method is based on determining the values of stresses in strictly designated points in front of the weld cap, depending on the

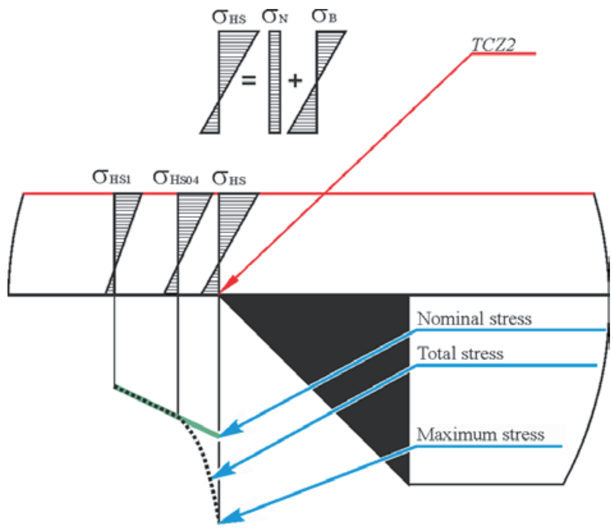


Fig. 8. Method of determining the nominal and maximum stresses at the front of the weld cap by using the Hot-Spot method and the interpretation of bending, normal and nominal stresses

thickness of the element, and calculating the hypothetical values of nominal stresses at the edge of the weld, in the location of the base of the concentrator. Next, the value of the coefficient of stress concentration is determined, based on the identified geometrical parameters of the weld R_s and α, and the maximum stresses are determined taking into account the nominal stresses calculated previously. In case of the existence of bending and stretching simultaneously, as is the case in the discussed example presented in Fig. 8, the coefficient of stress concentration should be determined independently from the bending and stretching. The values of reference stresses σ_{HS1} and σ_{HS04}, serving the purpose of calculating the nominal stresses σ_N and σ_G, can be acquired from the numerical solution *FEM*, for example in accordance with the procedures described in [14] and [16] or they can be determined experimentally with the strain gauge method.

The values of the stress concentration coefficients and maximum stress, for stretching and bending respectively, may be calculated from the equations published in [6, 10] or determined by numerical methods, such as conducting a separate calculation for plane problems. One of the methods of calculating the stress concentration coefficients, using nominal stresses in the Hot-Spot approach, was suggested by Monahan [10]. For pure stretching these coefficients' form is in accordance with the equation:

$$K_{t,hs}^{m} = 1 + 0,388^{\sim}\,0,37\left(\frac{t}{r}\right)^{0,454} \tag{1}$$

whereas for bending – the equation is:

$$K_{t,hs}^{m} = 1 + 0{,}512\cdot{}^{0{,}572}\left(\frac{t}{r}\right)^{0{,}469}. \tag{2}$$

The interpretation of the *t*, *r* and *Θ* parameters (in radians) is presented in Fig. 9. Other, more general equations, were suggested by Iida and Uemura [6]. These are less conservative [2] than the ones given by Monahan, but they consider the geometry of the welded joint to a greater extent, which is presented in Fig. 9. For stretching, the stress concentration coefficient is calculated in accordance to the equation:

$$K_{t,n}^{t} = K_{t,hs}^{m} = 1 + \frac{1-\exp\left(-0{,}9\cdot\sqrt{\frac{W}{2h}}\right)}{1-\exp\left(-0{,}45\cdot\sqrt{\frac{W}{2h}}\right)}\left[\frac{1}{2{,}8\left(\frac{W}{t}\right)-2}\frac{h}{r}\right]^{0{,}65} \tag{3}$$

and for pure bending – the formula is:

$$K_{t,n}^{b} = K_{t,hs}^{b} = 1 + \frac{1-\exp\left(-0{,}9\cdot\sqrt{\frac{W}{2h}}\right)}{1-\exp\left(-0{,}45\cdot\sqrt{\frac{W}{2h}}\right)}1{,}9\sqrt{\operatorname{tgh}\left(\frac{2t_p}{t+2h}+\frac{2r}{t}\right)}\operatorname{tgh}\left[\frac{\left(\frac{2h}{t}\right)^{0{,}25}}{1-\frac{r}{t}}\right]\frac{0{,}13+0{,}65\left(1-\frac{r}{t}\right)^{4}}{\left(\frac{r}{t}\right)^{1/3}} \tag{4}$$

where $W = 0{,}3(t+2h)(t_p+2h_p)$.

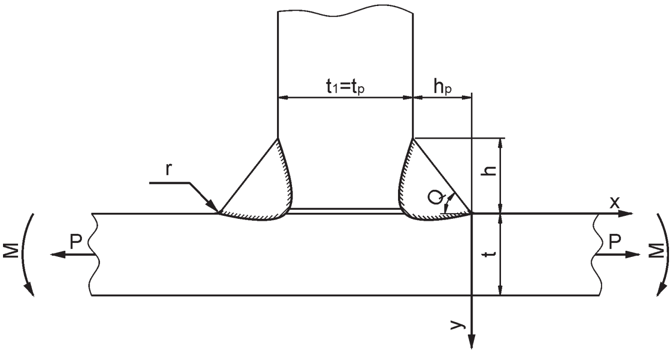


Fig. 9. Characteristic parameters of a welded joint [2] used in equations (1) – (4)

In the presented construction of a welded wheel, the third potentially critical spot was the sharp notch located in the area of the weld root, created after joining the disc and rim, marked in Fig. 4 as *TCZ3*. In this case, the durability criteria may be used based on fracture mechanics, which is to determine the values of the stress intensity factors *K* and their ranges Δ*K* in time of a full rotation of the wheel with applied radial force. The most useful tool allowing to determine the *K* value is the finite elements or the boundary elements method.

In the case of a wheel with a multi-arm disc, the area of rim damage is consistent with the location of *TCZ3*, which is located at the end of the weld, at the edge of the disc arm (Fig. 3).

4. Modification of the wheel construction

Based on the conducted model calculations and theoretical analyses, the construction of the welded wheel was changed in order to reduce the stresses in the critical zones. These modifications were to eliminate the areas of the highest concentration of stresses, located at the ends of the welds, by replacing the multi-arm disc with a full disc and the fragmented welds with a continuous circumference ones. In this way the concentration of stresses in the areas *TCZ2* and *TCZ3* was reduced, increasing the fatigue durability of the wheel, as shown by the results of later tests.

Fig. 10 presents the radial stress in the *TCZ1* zone during a full rotation of the modified wheel with a full disc around its axis, with applied force of *P* = 26 kN.

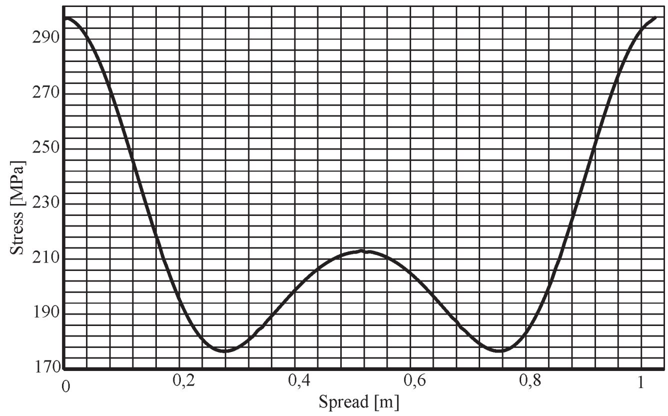


Fig. 10. Radial stress in the TCZ1 area during one full rotation of the wheel with a full disc 6mm thick, for force of P = 26 kN

As in the case of the wheel with a multi-arm disc, maximum stresses were occurring in the area closest to the appliance of radial force, and minimal values appeared in points corresponding to app. 1/3 of the wheel's circumference. In Fig. 10 two cycles of stress changes are seen during one full rotation of the wheel. A nearly 15% decline in the stress values was obtained relative to the wheels with a multi-arm disc, which is shown in Fig. 7.

The relation between the type of the wheel's construction and the value of weld parameters was also investigated. Collective results of local measurements of geometrical values of welds for different types of wheels are presented in Fig. 11. They point to the angle of the weld cap being independent from the wheel construction type, whereas the radiuses R_s are changing along with the thickness of the joined elements.

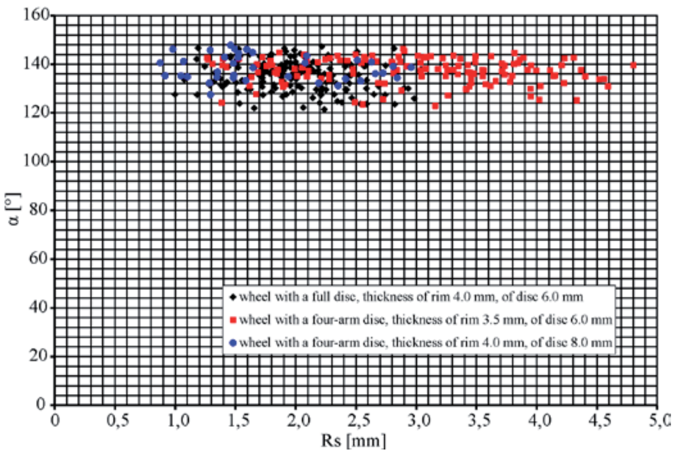


Fig. 11. The measured values of the weld cap angle in reference to the radius of weld/rim material border for different agricultural wheels construction solutions

The modified wheel constructions were tested on fatigue machines in a test for radial loads [18]. In Fig. 12a) fatigue crack is presented in the area of the rim bend, formed during a fatigue test of a wheel with a full disc. The origin location of the fatigue crack was located on the inner surface of the rim corresponding to the area of increased strains marked as *TCZ1* and identified with the *FEM*. A corresponding representation of the stresses calculated numerically is presented in Fig. 12b).

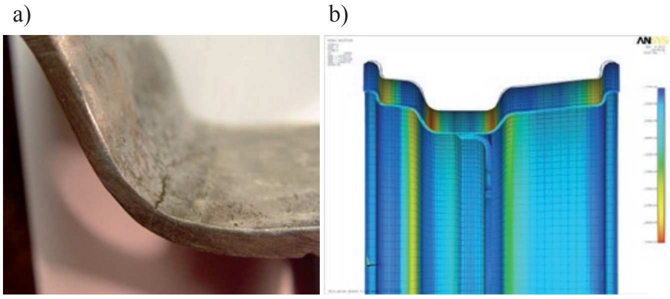


Fig. 12. a) The location of the fatigue crack in the bend of the rim, b) the corresponding distribution of stress in the FEM model, identified as TCZ1 (Fig. 4)

These wheels are characterised by significantly longer durability in comparison to the wheels with a multi-arm disc and a different location of fatigue cracks. The replacement of the multi-arm disc with a full disc and the fragmented weld with a continuous one allowed to eliminate local concentrations of stress at the ends of welds and remove the origin of fractures from the weld area. This changes also allowed to reduce the weight of the wheel itself by changing the thickness of the disc.

Fig. 13 presents the relation between durability and radial force for different types of wheels, obtained from fatigue tests [19]. The crosshatched area covers the location of the fatigue characteristics of wheels with multi-arm discs, whereas a continuous line marks the characteristic for a wheel with a full disc. A significant move to the right of the experimental curve is visible, which means an improvement in strength and increase in fatigue life. Another important feature of this new construction is the narrow spread of results for different levels of force applied, which allows to estimate more precisely the credibility range while predicting the durability of wheels.

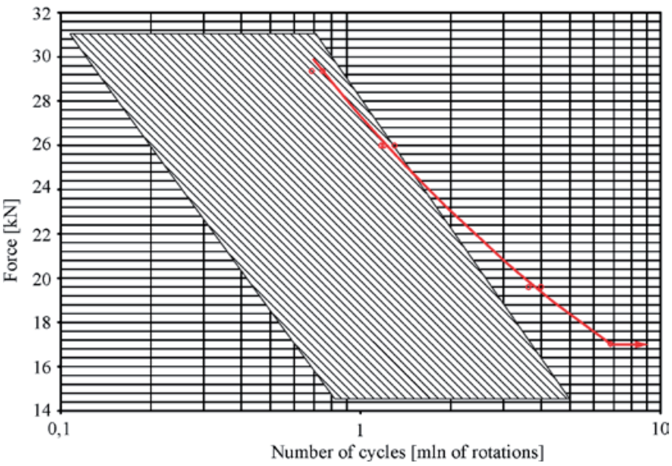


Fig. 13. A comparison of fatigue life of wheels with multi-arm discs with various disc and rim thickness (crosshatched area) and wheels with a full disc and circumferential weld (continuous line)

5. Procedure of fatigue designing of agricultural wheels

Fatigue design is connected to the shaping of the desired features of the product and therefore must involve many stages. The suggested algorithm of procedures in the case of wheels is shown in Fig. 14. The main elements here are the stages of *Assumptions* and *Prototypes,* which enable to achieve the desired goal.

The first stage of the process of shaping the utility features of wheels is the *Assumption Stage* which is defined after conducting the analysis of market requirements. In the case discussed here it is the need for designing a wheel of previously defined parameters, i.e.: weight, predicted durability, dimensions, method of mounting, construction type, type of tyre used, production cost also connected with the used technology etc. The process of creating the construction of designated durability is described in literature as *fatigue dimensioning of a product* [17].

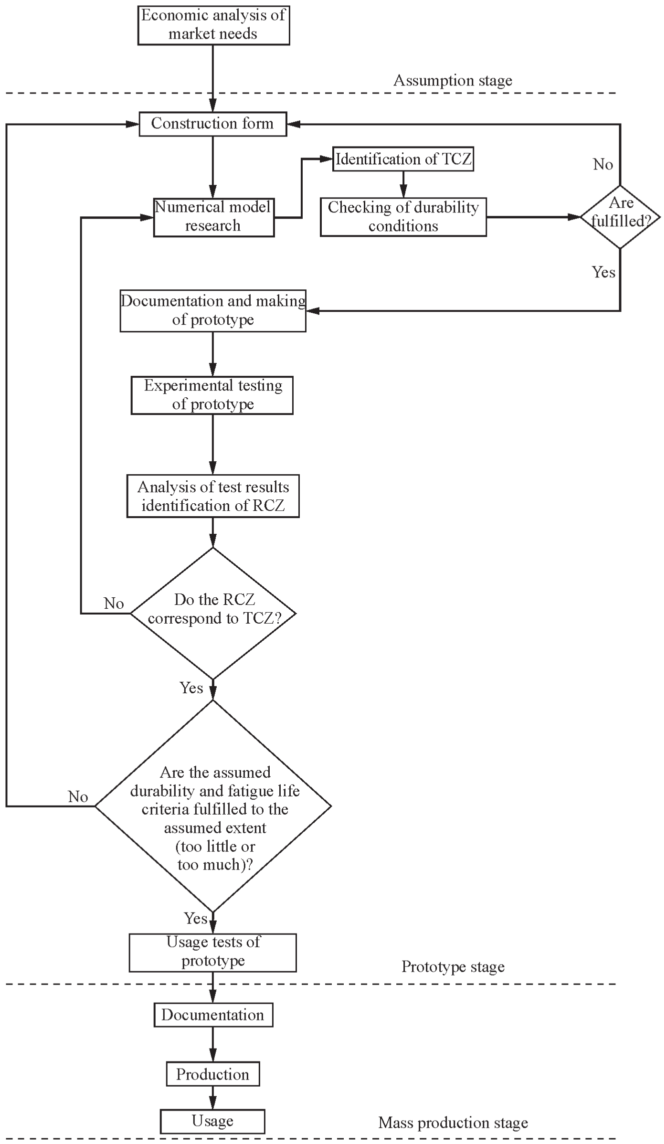


Fig. 14. The suggested procedure algorithm during fatigue design of welded wheels

The most developed, complex and time-consuming part of the algorithm is the *Prototype Stage,* which comprises of: conceptual work, numerical modelling aiming at identifying the areas of the wheel of greatest material stresses and the identification of *TCZ*, creating the prototype and testing it, which would verify the assumptions made, and any necessary modifications of the construction followed by testing, eventually leading to finding the optimal solution.

A key information obtained as a result of the conducted tests is determining the correspondence between the verified locations of fractures, i.e. the real critical zones *RCZ,* and the theoretical zones *TCZ*. If these areas overlap, the numerical model can be considered as qualitatively correct. However, in case of fractures appearing in places other than the theoretically estimated ones, the numerical model needs to be corrected for its accordance and precision. In the discussed case of the agricultural wheel the conformity of *TCZ* and *RCZ* was verified in experimental tests.

6. Summary and conclusions

The paper presented the method of *fatigue design* of a wheel used in agricultural vehicles, on the example of a welded wheel type 9.00x15.3. The local approach was used in fatigue design, where main effort has been made on the verification of zones of stress concentration and determining their local durability properties. The basis for theoretical analysis was the *FEM* model of a wheel with an *Implement* type tyre, which was positively verified with the use of the strain gauge method. As a result of the conducted numerical calculations three potential zones of an increased stress concentration were assessed, which determine the durability and fatigue life of the wheel. Their geometric qualities were identified and their local characteristic parameters were determined.

It was noticed that in welded wheels with multi-arm discs, the locations for the critical zones were the ends of the weld joints connecting the disc and rim, where a local increase in stresses occurred caused not only by the presence of the notch, but also by an uneven distribution of nominal stresses along the weld. These stresses initiated fatigue fractures which then propagated across the rim, causing the loss of the wheel integrity. It was also noticed that the value of the existing stress was determined by the thickness of the rim and that the thickness of the disc was of a smaller influence.

Based on the obtained results changes were made in the wheel construction, which were the replacement of a multi-arm disc with a full disc welded circumferentially with a decrease of its thickness, which in turn allowed to increase the fatigue life of the whole construction and lower its weight. Verification through experiment was conducted in a full scale test under radial loading. A significant quality improvement of the new construction, apart from the increase of fatigue life, was a high repetitiveness of the results of tests on every level of applied force, enabling more precise assessment of fatigue life of the structure.

References

1. Bureau Veritas. Fatigue strength of welded ship structures. Paris, 1994; BV NI393.
2. Chattopadhyay A. The GR3 Method for the Stress Analysis of Weldments. Waterloo. Ontario. Canada, 2009.
3. DIN 15018. Krane. Sthaltragwerke. Berechnunggrundsaltze, 1967.
4. E/ECE/324 E/ECE/TRANS/505. Uniform provisions concerning the approval of pneumatic tires for agricultural vehicles and their trailers. 2008.
5. Holscher H, Tewes M, Botkin N, Lohndorf M, Hoffmann, KH, Quandt E. Modeling of Pneumatic Tires by a Finite Element Model for the Development a Tire Friction Remote Sensor. Center of Advanced European Studies and Research. Ludwig-Erhard-Allee 2. Bonn. Germany, 2004; 53175.
6. Iida K, Uemura T. Stress concentration factors formulae widely used in Japan. Fatigue Fracture of Engineering Materials & Structures, 1996; 19(6): 779–86.
7. Jakubczak H. Niepewność danych w prognozowaniu trwałości zmęczeniowej konstrukcji nośnych maszyn. Zeszyt Mechanika. Warszawa: WPW, 2008; 194.
8. Mitas Agricultural Data book. Agricultural tires. Technical information 2007;. www.mitas.cz.
9. Małachowski J. Numerical study of tires behavior. Department of Mechanics and Applied Computer Science. Military University of Technology. Warsaw. Poland, 2007.
10. Monahan CC. Early Fatigue Cracks Growth at Welds. Computational Mechanics Publications. Southampton UK, 1995.
11. Mooney M. A theory for large elastic deformation. Journal of Applied Physics, 1940; 11, 582–597.
12. Oziemski S, Sobczykiewicz W. Konstrukcje nośne maszyn roboczych ciężkich. Podstawy teoretyczne i zasady projektowania. Warszawa: WPW, 1990.
13. PN-S-91240-03:1993. Koła z ogumieniem pneumatycznym wymagania i badania. PKNMiJ, Warszawa, 1993.
14. Poutiainen I., Tanskanen P., Marquis G. Finite element methods for structural hot spot stress determination – a comparison of procedures. International Journal of Fatigue, 2004, 26: 1147–1157.
15. Rivlin RS. Large elastics deformation of isotropic materials. VII. Experiments on the deformation of rubber. Philosophical Transactions of the Royal Society of London, 1951; Series A, 243: 251–288.
16. Savaidis G, Vormwald M. Hot-spot stress evaluation of fatigue in welded structural connections supported by finite element analysis, International Journal of Fatigue, 2000; 22: 85–91.
17. Sobczykiewicz W. Wymiarowanie w zakresie trwałości zmęczeniowej osprzętów MRC z uwzględnieniem procesu technologicznego wytwarzania i warunków eksploatacji. Prace Naukowe Instytutu Konstrukcji i Eksploatacji Maszyn Politechniki Wrocławskiej. Mechanika. Wrocław, 1987; 50: 267–276.
18. Tarasiuk P. Kształtowanie właściwości wytrzymałościowych kół pojazdów wolnobieżnych. Rozprawa doktorska. Oficyna Wydawnicza Politechniki Białostockiej, Białystok, 2010.
19. Tarasiuk P. Obliczanie mes kół pojazdów wolnobieżnych. Model opony a dokładność rozwiązania numerycznego. Acta mechanica et automatica. Politechnika Białostocka, Białystok, 2008; 06: 86–92.

Piotr TARASIUK, Ph.D. (Eng.)
Krzysztof Lech MOLSKI, Ph.D. (Eng.) Dr. Sc.
Andrzej SZYMANIUK
Bialystok University of Technology
ul. Wiejska nr 45C, 15-351 Białystok, Poland
E-mail: p.tarasiuk@pb.edu.pl, k.molski@pb.edu.pl, and.szy@gazeta.pl

Article citation info:
GRONOSTAJSKI Z, HAWRYLUK M, KRAWCZYK J, MARCINIAK M. Numerical modelling of the thermal fatigue of steel WCLV used for hot forging dies. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 129–133.

Zbigniew GRONOSTAJSKI
Marek HAWRYLUK
Jakub KRAWCZYK
Marcin MARCINIAK

NUMERICAL MODELLING OF THE THERMAL FATIGUE OF STEEL WCLV USED FOR HOT FORGING DIES

MODELOWANIE NUMERYCZNE ZMĘCZENIA CIEPLNEGO STALI WCLV STOSOWANEJ NA MATRYCE W PROCESIE KUCIA NA GORĄCO*

The paper presents an analysis of a numerical simulation of the low-cycle thermal fatigue of steel WCLV (X40CrMoV511) used in hot forging. As part of experimental studies a special test rig based on the rotating disc method was built and tests were carried out. Their results showed that the method can be used to reproduce the thermal fatigue conditions prevailing in the industrial forging process. For the given experimental conditions the instant when fatigue cracks appear was determined. A numerical model was built and the obtained finite element analysis results were compared with the laboratory test results in order to determine the amplitude of plastic strains at which a crack appears. As part of further research in the future the Coffin-Manson low-cycle fatigue model will be verified for other conditions and a low-cycle fatigue curve for steel WCLV will be determined.

Keywords: *low-cycle thermal fatigue, the rotating disc method,numerical modelling, die life, hot forging.*

W pracy przedstawiono analizę symulacji numerycznej niskocyklowego zmęczenia cieplnego stali WCLV stosowanej podczas kucia na gorąco. W ramach badań doświadczalnych zostało zbudowane specjalne stanowisko bazujące na metodzie "wirującego krążka" [13], przeprowadzone zostały próby, które potwierdziły możliwość stosowania tej metody do odwzorowania warunków zmęczenia cieplnego panujących w przemysłowym procesie kucia. Dla danych warunków eksperymentu określono moment pojawienia się pęknięć zmęczeniowych. Następnie zbudowano model numeryczny, po czym porównano uzyskane wyniki z MES i prób laboratoryjnych w celu określenia amplitudy odkształceń plastycznych, przy których pojawia się pęknięcie. Dalsze prace pozwolą w przyszłości na weryfikację niskocyklowego modelu zmęczenia Coffina-Mansona dla innych warunków oraz pozwolą stworzyć krzywą niskocyklowego zmęczenia stali WClV.

Słowa kluczowe: *niskocyklowe zmęczenie cieplne, metoda "wirującego krążka", modelowanie numeryczne, trwałość matryc, kucie matrycowe na gorąco.*

1. Introduction

Because of the extreme operating conditions the dies used in hot forging are exposed to many degrading phenomena. Forging dies are impacted by heavy mechanical and thermal loads and their surface is repeatedly intensively heated up, abraded and oxidized [8, 2, 3, 16]. Thermal fatigue is a major factor contributing to the wear of the tools. Thermal fatigue results from the large temperature gradient due to the changing contact with the preheated material. Because of the tool material's limited thermal conductivity there are large differences in temperature between the core and the surface. As a result, high stresses arise, especially in the die's surface layer. As the consequence of the cyclical variation in temperature the material is alternately tensioned and compressed. Also the dynamic loads exert a considerable impact, introducing additional stresses, which combined with the thermal stresses intensify fatigue leading to tool surface cracking. As a result, a network of fatigue cracks forms and an oxide film appears on the die's surface. As the number of forgings increases, a secondary network of cracks develops, the oxides separate from the surface and acting as an abrasive contribute to the abrasive wear of the tool. The network of cracks has an adverse effect on the quality of the finished product, by imprinting itself on its surface. A fatigue crack may also become the focus of a brittle fracture resulting in the total failure of the tool [5, 6, 1, 17]. Die life is a significant factor in the forging process costs. It is estimated that 10% of the price of a forging constitute tool expenses. Therefore research aimed at effective predicting, extending and optimizing the service life of forging tools is vital [14]. In recent years, not only ordinary nitriding, but also thermal endurance enhancing hybrid coatings have been increasingly often used [12, 10, 19, 7, 9]. Because of the large number and variety of factors having a bearing on forging tool life, this problem is very difficult to analyze. For this reason, increasingly often software tools, such as CAD, CAM and CAE, based on the finite element method (FEM), are employed to design, analyze and optimize forging processes [21, 15]. Information obtained from the FEM is very useful for building the fatigue models which can be used to predict the durability of tools in the real processes. It requires not only suitable experiments but also correct FEM models. There are a lot of research methods into thermal fatigue well described in literature where the thin-walled, discs and cylindrical samples heated by electric induction, the stream of the hot air etc. are applied [11, 13, 14, 19, 20]. Authors in position [4] describe the most popular research method of the thermal fatigue. The processes of thermal fatigue are very difficult to model by FEM regarding the rapid changes of the setting temperature. However, it permits to get the information indispensable to build the fatigue models.

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

2. Research aim and range

The aim of this research was to build an accurate numerical model of the rotating disc method, enabling one to determine the amplitude of the plastic strain resulting from a temperature gradient. As part of the further research the Coffin-Manson low-cycle fatigue will be verified on the basis of the number of cycles after which first cracks appear at the strain amplitude determined for the given conditions by FEM. The Coffina-Manson fatigue model should be correct for thermal and mechanical loads because both kinds of loads result in deformations, which lead to destruction of material.

3. Description of industrial forging of spur gear

The considered industrial process of forging the spur gear in the Jawor Forge PLC consists of three operations performed on a crank press with a load of 25MN (Fig. 1). The first forging operation consists in the upsetting of the cylindrical preform. The second operation is die roughing. In this operation that the greatest forging pressures and material deformations occur. In the third operation, i.e. finishing die forging, the forging assumes a shape close to that of the finished product. A detailed analysis of the process showed that abrasive wear, thermal fatigue and thermal-mechanical fatigue are the dominant material degrading phenomena. Figure 2 shows photographs of the die inserts after a different number of forgings, in which the impacts of the above phenomena are visible.

Fig. 1. Die inserts used in industrial hot forging.

a) b) c)

Fig. 2. Die insert wear after forging: a) 550, b) 1900 and c) 4300 pieces.

4. Thermal fatigue investigations by rotating disc method

In real conditions it is often impossible (and if possible, it takes much time and money) to test the thermal endurance of a material. Then numerical simulations can be run, but they do not always accurately represent the operating conditions of the components or the behaviour of the material. A special test rig is the right tool for the physical modelling of this phenomenon. Having carried out a survey of the literature [5, 11, 13, 20] on the resistance to high temperature gradients and on the peculiarities of the spur gear forging process, the authors designed and built a test rig based on the rotating disc method [18], as well as proposed a test methodology and carried out tests. The arguments for the choice of this method were that thermal fatigue tests are carried out on geometrically simple specimens in easily controllable conditions and that a numerical model reproducing the experiment can be built. The idea of the investigations consists in combining the physical and numerical modelling of the thermal fatigue phenomenon. During testing the material is heated up in its surface layer, similarly as in the hot forging process. A specimen, in the shape of a disc with a hole, was made from tool steel WCLV (the die material). It was subjected to the same thermal treatment as the forging tools, i.e. quenching and double tempering at a temperature of 540°C. The test rig is shown in fig. 3. The disc (8), immersed in water flowing through a tank (8), rotates at constant speed and is superficially (to a depth of about 1-2 mm) heated up by a high-frequency inductor (6) with temperature smoothly adjustable in a range of 100-900°C.

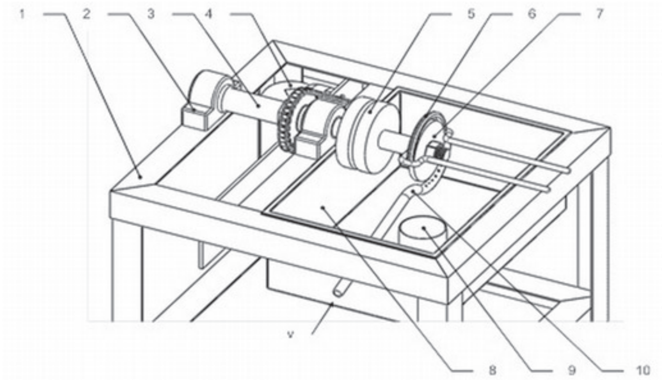


Fig. 3. Scheme of thermal fatigue test rig; 1- steel frame, 2- support with plain bearing, 3- shaft drive, 4- DC motor 5- cooling rings, 6- inductor, 7 - test specimen, 8 - water tank, 9 - water outflow, 10 - adjustable water supply.

In the course of heating the near-surface layer locally very quickly expands while during rapid cooling it undergoes rapid compression.

Figure 4 shows a photograph of the industrial forging process with temperature distributions on the particular dies, taken by a thermovision camera. On this basis proper thermal parameters (the cycle upper temperature T_g=650÷800°C), corresponding to the real process, were selected. Figure 5 shows a thermogram with temperature distributions on the tested specimen. The thermogram was the basis for building a numerical model. Two cycle lengths: 15 s – the actual cycle and 4 s – the shortened cycle were used in the investigations. To make the analysis complete, the average maximum depth of cracks, the density of the cracks and their maximum depth were adopted as the criteria for evaluating the fatigue resistance of the material. Temperature was measured by a thermovision camera and a pyrometer. The cycle length was determined by the motor's rotational speed which could be changed by changing the current intensity and voltage.

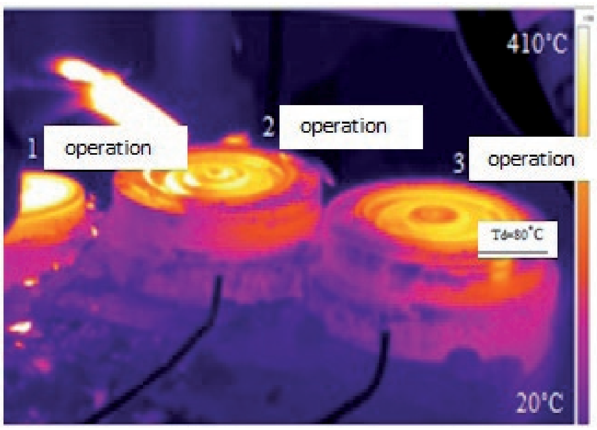


Fig. 4. Thermogram of surfaces of die inserts.

A decision was made to carry out several tests for different cycle lengths and different cycle upper temperatures. In the course of the test the condition of the surface was being documented by a camera. It was critical to determine the instant at which the first cracks appear on

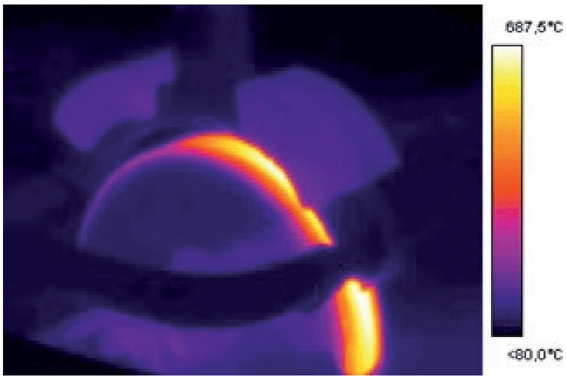


Fig. 5. Thermogram of specimen surface.

the surface of the specimen (observation under optical microscope). Figure 6 shows a typical network of cracks on the specimen's surface at temperature Tg=700°C at which fine cracks appeared already after 50 fatigue cycles. The direction of the cracks was consistent with that of the specimen axis from the middle of the heated path towards the edge. The network had an open character and it expanded with the number of cycles.

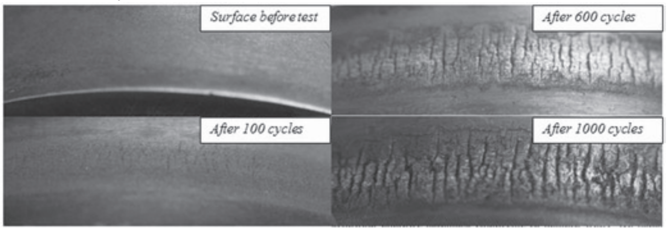


Fig. 6. Network of cracks after successive thermal fatigue cycles in test conditions: cycle length- 5 s, upper cycle temperature 700°C.

5. Numerical modelling

A numerical model was built for the test carried out at T_g=700°C for 5 s. Fifty cycles were simulated since after this number of cycles cracks would appear during the physical test. Rotating disc method simulations were run using the MARC software package. A full (3D) model of the process (fig. 7) was built and the symmetry plane was used to shorten the computing time. The model is made up of 28590 hexa elements. Less thick elements and a denser mesh were used on the face of the disc in order to more accurately reproduce plastic deformations. In the model the heating part (simulating induction heating) and the cooling part (simulating cooling in water) rotate around the stationary disc. The heating part supplies thermal energy to the disc while the cooling part removes this energy by means of the NEAR CONTACT function. The complete rotation takes about 5 seconds (200 simulation steps). The material specifications were taken from the MATYLDA material database and from dialatometric studies. 20°C and 50°C were adopted as respectively the ambient temperature and the initial temperature of the disc.The temperature of the heating part and that of the cooling part was respectively 1450°C and 20°C. The thermal conductivity coefficients in the contact between the heating and cooling parts and the disc were assumed to be respectively 8 kW/m²/K and 9 kW/m²/K. The thermal conductivity coefficient (selected from material databases) between the disc and the surroundings amounted to 0.35 kW/m²/K.

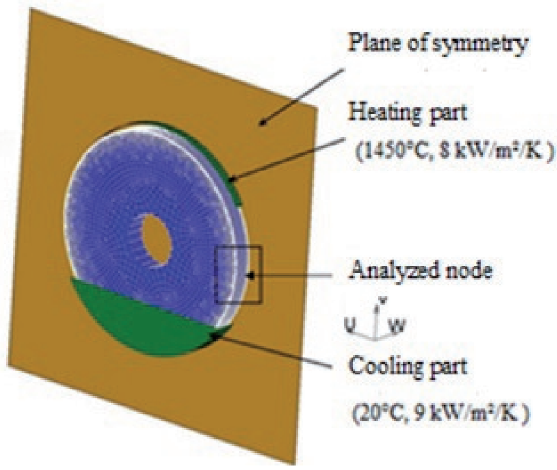


Fig. 7. General view of model

6. Simulation results

Figure 8 shows the temperature distribution after 50 fatigue cycles (1 cycle=5s) in the symmetry plane where the largest variation in temperature occurred. A diagram of temperature variation in the course of 50 cycles for a selected node is shown in Fig. 9. The results are consistent with the laboratory test results as regards both the temperature variation and distribution (as revealed by the thermogram) on the surface of the specimen.

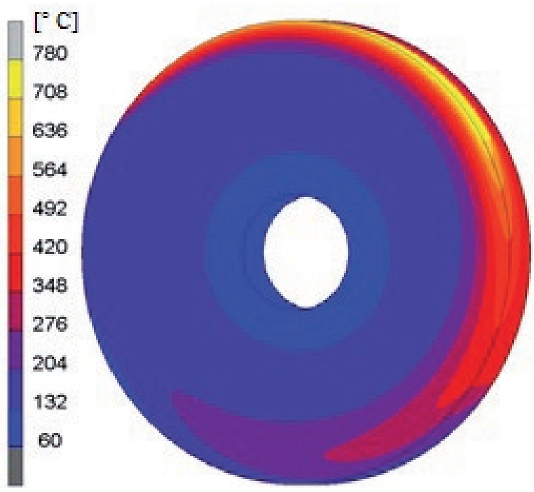


Fig. 8. Temperature distribution after50 cycles.

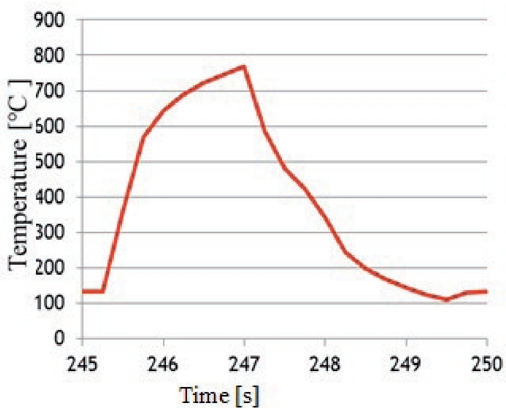


Fig. 9. Distribution of temperature over time (50th cycle).Temperature

Figure10 shows distributions of circumferential strain after 50 cycles since in the case of this method it is these strain components which are mainly responsible for the development of cracks on the side surface of the tested specimen. In the close-up one can see the distribution of strain in the symmetry plane – the place where the

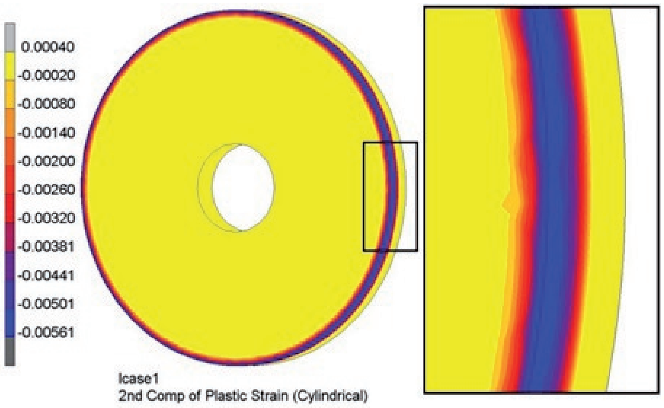


Fig. 10. Distribution of circumferential strain E2 after 50 cycles.

highest strains occur. During the experiments the initiation and propagation of cracks were observed in the specimen surface directions. Figure 11 shows a diagram of circumferential strain over 50 fatigue cycles for a selected node. Positive and negative changes in strain in the course of the whole process are visible. The decrease in circumferential strain amplitude is due to the strain hardening of the material

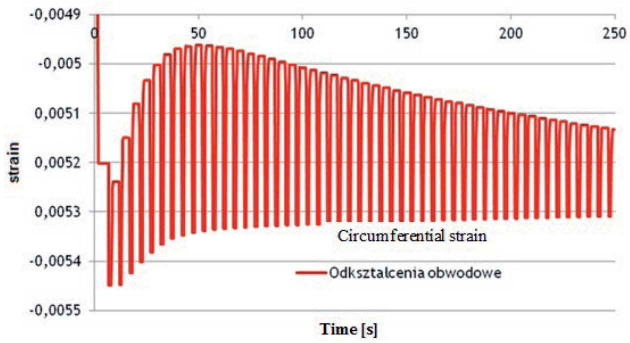


Fig. 11. Circumferential strain-time diagram for selected node.(Strain)

in the successive cycles.

Then the average strain amplitudes and the instant at which cracks appear after a given number of cycles were determined and it will be used to verify the Coffin-Manson low-cycle fatigue model (1).

$$\frac{\Delta\varepsilon_p}{2} = \varepsilon_f'(2N)^c \tag{1}$$

where:

- $\Delta\varepsilon_p/2$ – the plastic strain amplitude,
- ε_f' – a coefficient of fatigue ductility
- N – the number of cycles until fatigue,
- c – a fatigue ductility exponent.

ε_f' is approximately equal to the real elongation under monotonic tension and c for most materials changes in a range from −0.7 to −0.5.

Then by changing the test conditions (T_g or the cycle length) and determining the instant when a crack appears during the rotating disc fatigue test one can obtain another circumferential plastic strain amplitude and thereby determine the next points on the low-cycle thermal fatigue characteristic curve. For the diagram obtained in this way it is possible to determine the limit number of cycles for the given tool material in the industrial forging process, after which one can expect fatigue cracks to appear on the surface of the dies. In the case of forging processes, one should also take into account the possibility of cracking as a result of mechanical deformations.

7. Conclusion

The authors using a test rig based on the rotating disc method have created a numerical model of the rotating disc method. Physical tests were carried out to determine the instant when a crack would appear in the specimen's side surface in the given test conditions. The numerical modelling highly accurately reproduced the rotating disc method, and so the course of the real process (but with regard to only thermal loads). In order to reveal very small plastic strains, a denser mesh was applied to the disc face (the finite element thickness was reduced to about 0.1 mm). Thanks to the adopted boundary conditions the stresses and strains occurring in the disc in the course of the process, which could not be physically measured,were determined. The amplitude of the plastic strains needed to determine the limit number of cycles until a crack appears was determined for the forging tools through numerical modelling.

***Acknowledgement**:*

This research was funded as part of the authors' own project no.N508 476038.

Literature

1. Ebara R. Fatigue crack initiation and propagation behavior of forging die steels. International Journal of Fatigue 2010; 32: 830–840.
2. Gronostajski Z, Hawryluk M, Kaszuba M, Sadowski, Walczak S, Jabłoński D. Measuring &control system in industrial die forming processes. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2011; 3: 62–69.
3. Gronostajski Z, Hawryluk M, Kaszuba M, Polak S, Niechajowicz A, Walczak S, Jabłoński D. Die Profile Optimization for forging constant velocity joint casings. Archives of Metallurgy and Materials 2011; 56(2): 551–558.
4. Gronostajski Z, Hawryluk M, Zwierzchowski M, Kaszuba M, Marciniak M. Analiza zmęczenia cieplnego stali wclv stosowanej na matryce do kucia na gorąco. Rudy i metale nieżelazne 2011; 56(11):654–660.
5. Grüning A, Lebsanft M, Scholtes B. Cyclic stress–strain behavior and damage of tool steel AISI H11 under isothermal and thermal fatigue conditions. Materials Science & Engineering 2010; 527(7-8): 1979–1985.
6. Katunin A. The conception of the fatigue model for layered composites considering thermal effects. Archives of Civil and Mechanical Engineering 2011; 11(2): 333–342.
7. Kot M. Contact mechanics of coating-substrate systems: Monolayer and multilayer coatings. Archives of Civil and Mechanical Engineering 2012; 12(4): 464–470.
8. Lavtar L, Muhic T, Kugler G, Tercelj M.Analysis of the main types of damage on a pair of industrial dies for hot forging car steering mechanisms. Review article. Engineering Failure Analysis 2012; 18(4): 1143–1152
9. Mrva P, Kottfer D, Kaczmarek Ł. Effect of shot peening and NiAl coating on fatigue limit of Mg-Al-Zn-Mn Alloy. Archives of Metallurgy and Materials 2011; 56(3): 743–748.
10. Paschke H, Weber M, Yilkiran T. Optimized Plasma Nitriding Processes for Efficient Wear Reduction of Forging Dies. Archives of Civil and Mechanical Engineering 2012; 12(4): 407–412.

11. Sjöström J, Bergström J. Thermal fatigue testing of chromium martensitic hot-work tool steel after different austenitizing treatments. Journal of Materials Processing 2004; 153-154: 1089–1096.
12. Smolik J. Hybrid Surface Treatment Technology for Increase of Hot Forging Dies. Archives of Metallurgy and Materials 2012; 57(3): 657–664.
13. Starling C, Branco J. Thermal fatigue of hot work tool steel with hard coatings. Thin Solid Films 1997; 308-309: 436–442.
14. Summerville E, Venkatesan, K, Subramanian C. Wear processes in hot forging press tools. Materials and Design 1995; 16(5): 289–294.
15. Szeliga D, Kuziak R, Pidvysotskyy V, Pietrzyk M. Rheological model of Cu based alloys accounting for the preheating prior to deformation. Archives of Civil and Mechanical Engineering 2011; 11(2): 451–467.
16. Świątoniowski A, Sińczak J, Łukaszek-Sołek A, Schmidt J. Analysis of Forging Process of the NiCrN Superalloy for Motor Boat Driving Shaft 2012; 57(3): 719–725.
17. Tomaszek H, Jasztal M, Zieja M. A simplified method to assess fatigue life of selected structural components of an aircraft for a variable load spectrum. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2011; 4: 29–34.
18. Weroński A. Zmęczenie cieplne metali. Warszawa: Wydawnictwa Naukowo-Techniczne, 1983.
19. Yilkiran T, Behrens B.-A, Paschke H, Weber M, Brand H. The potential of plasma deposition techniques in the application field of forging processes. Archives of Civil and Mechanical Engineering 2012; 12(3): 284–291.
20. Yong-an M, Xiao-chun W, Luo-ping X, Wen-jun T, et. al. Influence of different surface treatments of H13 hot work die steel on its thermal fatigue behaviors. Journal of Shanghai University 2001; 5(4): 326 – 330.
21. Yu TT, Liu P. Improved implementation of the extended finite element method for stress analysis around cracks. Archives of Civil and Mechanical Engineering 2011; 11(3): 787–805.

Prof. Zbigniew GRONOSTAJSKI, Ph.D. hab. (Eng.)
Marek HAWRYLUK, Ph.D. (Eng.)
Jakub KRAWCZYK, M.Sc.
Marcin MARCINIAK, M.Sc.
Institute of Production Engineering and Automation
ul. Ignacego Łukasiewicza 5, 50-371, Wrocław, Poland
e-mail: Zbigniew.Gronostajski@pwr.wroc.pl

Article citation info:
PARCZEWSKI K. Effect of tyre inflation preassure on the vehicle dynamics during braking manouvre. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 134–139.

Krzysztof PARCZEWSKI

EFFECT OF TYRE INFLATION PREASSURE ON THE VEHICLE DYNAMICS DURING BRAKING MANOUVRE

WPŁYW CIŚNIENIA W OGUMIENIU NA DYNAMIKĘ RUCHU POJAZDU PODCZAS MANEWRU HAMOWANIA*

The paper presents the problem of reducing the impact of inflation pressure on the tires, and the weight distribution of the vehicle on the road. The results presented in this publication are based on the research bench and passenger vehicle equipped with anti-lock braking system. Bench testing was conducted stiffness of tires and road tests, performing maneuvers based on ISO standards, braking in the straight patch of road and of the twisting road.

Keywords*: vehicle dynamics, tyre, vehicle testing, braking, wheel sideslip angle, vehicle stability.*

W publikacji przedstawiono zagadnienie wpływu obniżenia ciśnienia w oponie na charakterystykę opon, rozkład nacisków i zachowanie się pojazdu na drodze. Wyniki prezentowane w publikacji oparto na badaniach stanowiskowych i drogowych samochodu osobowego wyposażonego w układ zapobiegający blokowaniu kół. Przeprowadzono badania stanowiskowe sztywności opon oraz badania drogowe, polegające na wykonywaniu manewrów opartych na normach ISO: hamowania na prostoliniowym odcinku oraz na łuku drogi.

Słowa kluczowe*: dynamika ruchu, opony, badania pojazdu, hamowanie pojazdu, kąty znoszenia kół, stabilność pojazdu.*

1. Introduction

During the vehicle motion for the transmission of power from the vehicle are the responsibility of the state of the tire and the road surface. The forces transmitted to the road depend on the parameters of the vehicle and its movement. The uneven tire pressure changes the stiffness of the tires both radial and longitudinal and transverse, as well as changing wheel rolling and load resistance provided by the vehicle to the ground. It should also be noted that the phenomenon of the contact patch with the road affect the stability and controllability of the vehicle. Analysis of these phenomena has been carried out and the static conditions - when measured as a tire and vehicle motion conditions on the test track.

2. The tire inflation pressure characteristics

The study of the influence of tire inflation pressure on their characteristics dealt with especially in laboratories engaged in researching tires. The first comprehensive work on the tires was presented by S. Clark and others [4]. One of the most important laboratories studying tire center in Delft (TNO Automotive - Netherlands). The center of this is related to HB Pacejka dealing with modeling of tires. Published several works related to the study and modeling of tires, including tire cooperation model was established with the road, known as MF (magic formula) [11]. The issue of inflation pressure in the tire shown in the literature [3, 10, 15]. Similar tires studies were conducted in different laboratories, including lab CRREL (U.S. Army Cold Regions Research and Engineering Laboratory in Anchorage) [8].

Posted in literature results allow to conclude that the reduction in tire inflation pressure reduces directional and angular stiffness of tires. To the relation of directional and angular tire stiffness where substituted individual forces and moments from the model MF, the index *nom* refers to the tire stiffness in the nominal inflation pressure.

Radial stiffness was determined from the formula [3]:

$$C_{Fz} = \left.\frac{dF_z}{d\rho_z}\right|_{\rho_z=0} = (1+q_{CFz3}\cdot dp_i)\cdot C_{Fz,nom} \tag{1}$$

where ρ_z – radial deflection of tires, q_{CFzi} – MF tire model parameter, $C_{Fz,nom}$ – radial stiffness at nominal inflation pressure in the tires, dp_i – coefficient of pressure change, where p_i is the real pressure and p_{i0} is the nominal pressure.

$$dp_i = (p_i - p_{i0})/p_{i0} \tag{2}$$

Longitudinal stiffness was determined from the following formula [3] assuming a constant loading force F_z:

$$C_{Fx} = \left.\frac{\partial F_{xW}}{\partial d_x}\right|_{d_x=0} \tag{3}$$

$$C_{Fx} = C_{Fx,nom}\cdot(1+q_{CFx3}\cdot dp_i + q_{CFx4}\cdot dp_i^2) \tag{4}$$

where: d_z – longitudinal deflection of the tire, q_{CFxi} – MF tire model parameters, $C_{Fx,nom}$ –longitudinal stiffness of the tire at nominal inflation pressure.

Lateral stiffness was determined as above [3] assuming a constant loading force F_z:

$$C_{Fy} = \left.\frac{\partial F_{yW}}{\partial d_y}\right|_{d_y=0} \tag{5}$$

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

$$C_{Fy} = C_{Fy,nom} \cdot (1 + q_{CFy3} \cdot dp_i) \quad (6)$$

where: d_y – lateral deflection of the tire, q_{CFyi} – MF tire model parameters, $C_{Fy,nom}$ – lateral stiffness of the tire at nominal inflation pressure.

Tire torsional stiffness was determined as above [3] assuming a constant loading force F_z:

$$C_{Mz} = \left.\frac{\partial M_{zW}}{\partial \psi}\right|_{\psi=0} \quad (7)$$

$$C_{Mz} = C_{Mz,nom} \cdot (1 + q_{CMz1} \cdot dp_i) \quad (8)$$

where y – angular deflection of the tire, q_{CMi} – MF tire model parameters, $C_{Mt,nom}$ – torsional tire stiffness at nominal inflation pressure.

As can be seen in all relationships there is an additional factor $(1+q_{MFi} \cdot dp_i)$, which is the product of the inflation pressure change coefficient dp_i and MF tire model parameter q_{MFi} except longitudinal stiffness in the case where the additional element is described in a quadratic function.

Similarly, reduction the inflation pressure in the tire increases the rolling resistance. In the literature [8] was described, the empirical relationship shown below:

$$\mu_r = \frac{K}{1000} \cdot \left(5.1 + \frac{5.5 \cdot 10^5 + 90 \cdot F_z}{p} + \frac{1100 + 0.0388 \cdot F_z}{p} \cdot v_x^2 \right) \quad (9)$$

where: K – coefficient depending on the tire construction (0.8 for radial tires, 1.0 for diagonal tires), v_x - tire longitudinal velocity [m/s], p – inflation pressure [MPa].

Dissipated power to overcome the rolling resistance will be so dependent on the pressure in the following way:

$$P = F_r \cdot v_x = -\mu_r \cdot v_x \cdot F_z =$$
$$= \frac{-K \cdot v_x}{1000} \cdot \left(5.1 + \frac{5.5 \cdot 10^5 + 90 \cdot F_z}{p} + \frac{1100 + 0.0388 \cdot F_z}{p} \cdot v_x^2 \right) \cdot F_z \quad (10)$$

where: P – power lost in overcoming rolling resistance, F_r – rolling resistance force.

3. Investigation of tire characteristics

The tire stiffness test for stationary conditions was carried out in the Laboratory of Vehicles, University of Bielsko-Biala. For the measurement was selected tires of size 155/60R14 with nominal inflation pressure 0.22 MPa. The tire was placed on the moveable base during tests. Vertical force was exerted on the axis of the wheel (in the case of measuring the radial stiffness, the force is changed from 0 to a maximum of ~ 50% higher than the static load), and in other cases (measured in longitudinal and lateral stiffness) tires were tested at two values of the loading force. Tire deflection measurement was performed using an optical displacement meter. Tangential force was applied to the moveable base, on which was a wheel. Measurements were carried out at different values of tires inflation pressure of minimum 0.06 to 0.3 MPa. Force measurements were carried out using strain gauges.

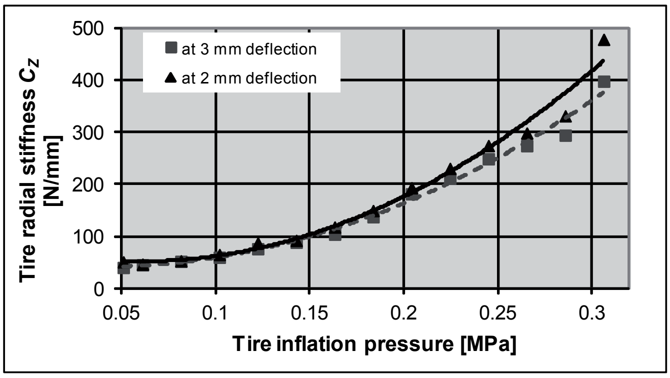


Fig. 1. The changes of tire radial stiffness C_Z as a function of inflation pressure to deflection 2 and 3 mm (own research)

The research obtained the radial, longitudinal and lateral stiffness characteristics as a function of the tire inflation pressure and the tire model parameters q_{MFi} (Figure 1-3).

The measurements suggest that the radial stiffness increases with increasing pressure and the nominal inflation pressure (~ 0.22 MPa) stiffness is C_z =~ 200 N/mm. The parameter model MF, $q_{CFz3} = -0.550$ for the pressure p = 0.1 MPa.

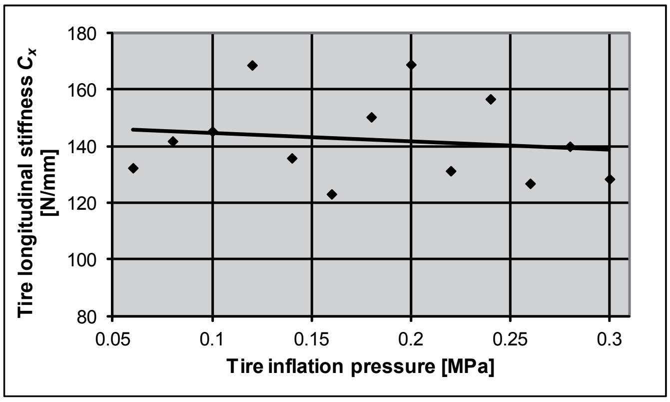


Fig. 2. The changes of tire longitudinal stiffness C_X as a function of inflation pressure (own research)

The measurements suggest that the change in the longitudinal stiffness of the tire to a slightly dependent on the inflation pressure in the tire, and its value is maintained at a constant level with a slight decline C_X = ~ 145 N/mm. The parameter model MF, q_{CFx2} = 0.035 for the pressure p = 0.1 MPa.

Lateral stiffness of the tire in the inflation pressure range of 0.22 to 0.27 MPa is the highest and is about 105 to 115 N/mm. The parameter model MF, $q_{CFy3} = -0.427$ for the pressure p = 0.1 MPa.

Tire stiffness parameters affect both the behavior of the vehicle at the overcoming rough roads and during turning maneuvers. In addition, the radial stiffness of the tire influence on the deflection of the suspension (as shown in the following part).

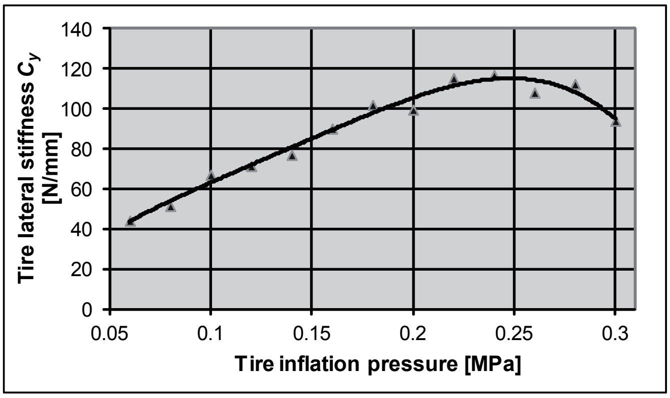


Fig. 3. The changes of tire lateral stiffness C_Y as a function of inflation pressure (own research)

Figure 4 shows the percentage change in reaction force volume of the suspension (in a quasi-static conditions) depending on the tire inflation pressure and the load operate to the wheel.

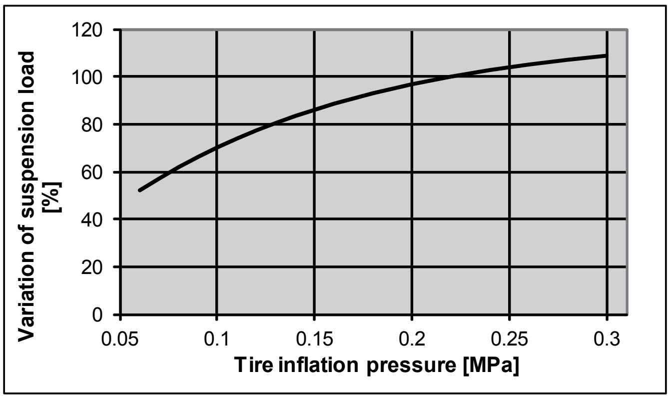


Fig. 4. The percentage changes of suspension load F_Z as effect of tire inflation pressure

At low inflation pressures, lower by 0.1 MPa from the nominal pressure, the reaction force in the suspension is reduced by ~ 30%, the increase in tire pressure of 0.1 MPa above the nominal pressure, increases the reaction force of ~ 10%.

4. The behavior of tire inflation pressure affect on vehicle motion

4.1. Assumptions for the road test

Vehicle Division of University of Bielsko-Biala, led the research vehicle based on ISO standards [13, 14]. For comparisons of selected two trials: braking on straight and curvilinear section of road. In both studies the impact of the driver on the test results is relatively small and the tire impact is clearly noticeable. The force on the brake pedal was large enough to activate the ABS system [5, 7].

For safety reasons, the test was performed on clean dry asphalt. For testing uses several types of sensors: pressure sensors mounted in the brake system, the force sensor on the brake pedal, sensors to measure the longitudinal and transverse velocity, acceleration sensors allowing to measure in the 3 direction, sensors to measure the angular velocity of the car body, sensors to measure the angle of rotation and torque on the steering wheel [12, 13]. Vehicle weight was due to its own weight, driver and measuring equipment.

During braking, the wheels normal forces are changing with deceleration. This changes the limits of adhesion forces and as a result of the launch of anti-lock braking system (ABS) and reduction of braking forces generated by individual wheels brakes. Load of each wheel was determined by measuring the center of gravity position, and the longitudinal and lateral forces arising from the mot ion conditions. In determining the normal forces on individual wheels does not include the impact of the stabilizer. Based on the measured pressure in the brake system and the brakes set of geometrical parameters for braking force at individual wheels.

Presents the results of two tests – test vehicle braking maneuver performing at a straight and at the curvilinear section of the road.

4.2. Braking test of the straight section of road

The first test was carried out on a straight section of road. The driver continued straight direction. After obtaining the appropriate speed pressed for the brake pedal. Braking force was so large that it has launched the ABS system.

Braking test on straight section of road shows (Fig. 6) that the reduced inflation pressure in one of the car wheels, equipped with ABS,

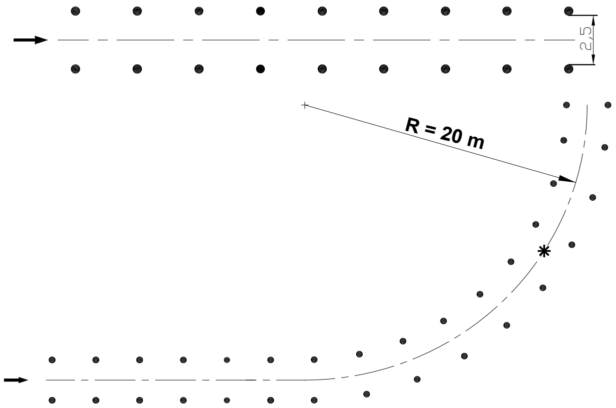


Fig. 5. Vehicle trajectories of individual tests: a) braking of the straight section of road,b) braking of the curvilinear section of road

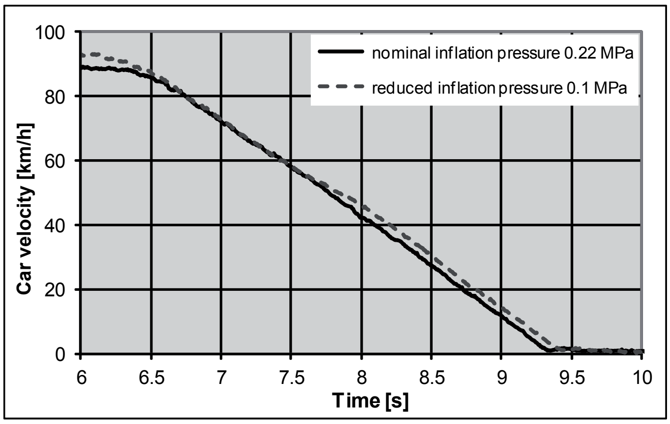
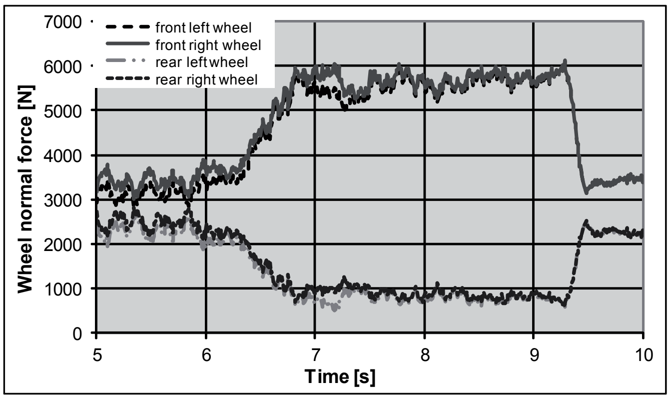


Fig. 6. The change of the vehicle velocity during brake test on a straight section of the road (the inflation pressure in the right front tire was nominal and reduced)

a)



b)

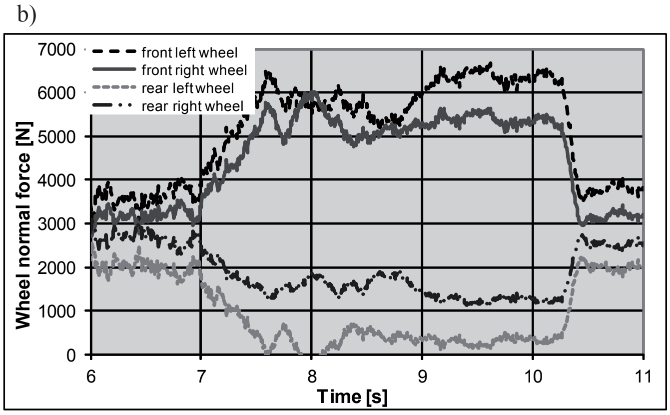
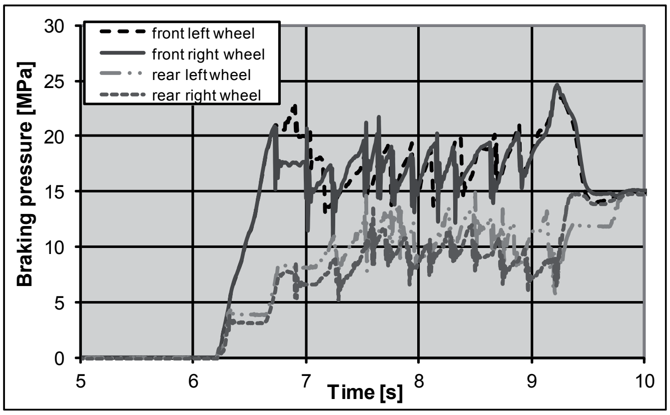


Fig. 7. Wheel load of the vehicle during test to brake on a straight section of the road (the inflation pressure in the right front tire was nominal (a) and reduced (b))

will result in a smaller change in velocity and an increase in vehicle stopping distance. It should be noted that braking test was carried out on a clean, dry road, with a relatively high adhesion. If the test performed in worse terms of grip, speed reduction would be smaller and the difference of stopping distances would increase.

The diagram shows the difference of normal forces on individual wheels acting on road during deceleration (Fig. 7). Reduction of the inflation pressure in one of the wheels to change the distribution of loads and affects their activities of the ABS - as shown in the diagrams in the pressure at brake circuits of individual wheels.

a)



b)

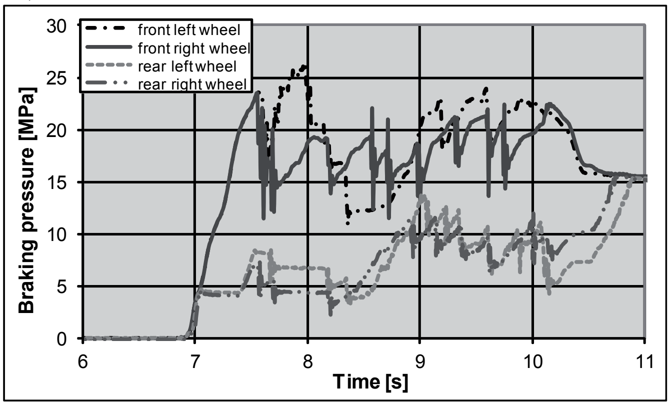


Fig. 8. Pressure in the brake circuit acting on each wheel (right front tire inflation pressure was nominal (a) and reduced (b))

In Figure 8, for the braking of reduced inflation pressure in the right front wheel, are shown in greater braking pressure correction (leading to a reduction in braking force) in the individual circuit of the wheel, and much more time offset to ensure straight track.

4.3. Braking test of the curvilinear section of road

The second test was carried out on the curvilinear section of road. The driver held the steering wheel in such a way that the vehicle is moving along a circular path. After defeating about 15 m pushing on the brake pedal. Pedal force provided a launch of the ABS system. Figure 9 shows the paths of the vehicle during the test. Driving tracks of both vehicles are similar despite considerable differences in the angle of the steering wheel rotation. For a similar trajectory, the angle of the steering wheel of a vehicle with a reduced inflation pressure in one of the wheels, was greater by about 35 degrees.

On the following graphs shows the waveforms pressure in the brake system acting on the brakes at individual wheels of the vehicle (Fig. 10) as well as waveforms changes in wheels normal forces performed under the turning maneuvers and the braking (Fig. 11).

Figure 10 shows the apparent size differences pressures (and consequently braking forces) front and rear axles, corrected for the nor-

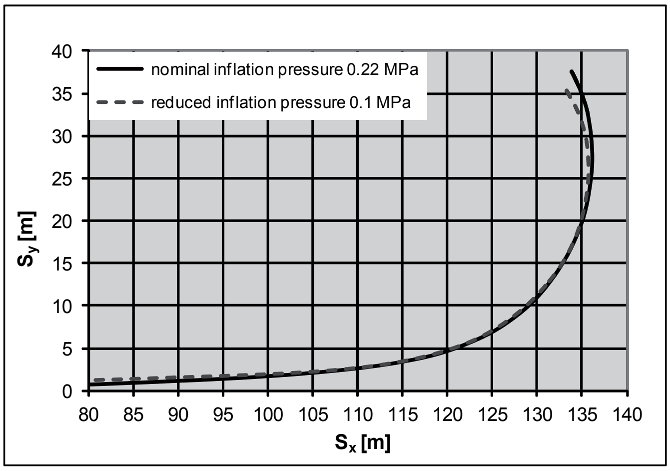
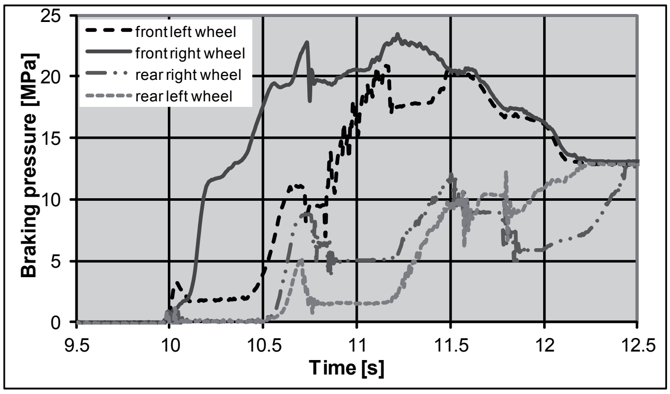


Fig. 9. The path of the vehicle during test to brake in a curvilinear section of road

a)



b)

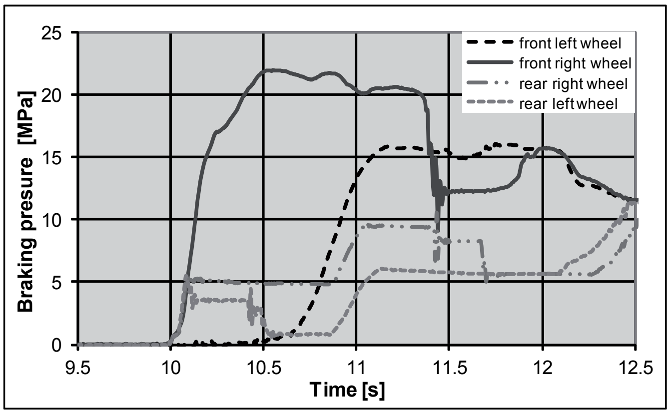


Fig. 10. Pressure in the brake circuit acting on each wheel (right front tire inflation pressure was nominal (a) and reduced (b))

mal force distribution and the centripetal force acting on the vehicle during braking on curved track. It can be seen that in the initial phase of braking the rear left wheel normal force is close to zero, resulting in a reduction of pressure, by the system ABS in the rear wheel brake circuit, and thus decrease the braking forces to small values. In the case of braking with the lower inflation pressure in the front right wheel, there is a significant adjustment braking pressures acting on the front left wheel and increasing the pressure in the rear axle braking circuit. The pressure in the front brake circuit is greatly increased in the first stage of braking and then rapidly decreases in the second stage of braking. There is a marked increase in the pressure in the rear wheel brake circuits.

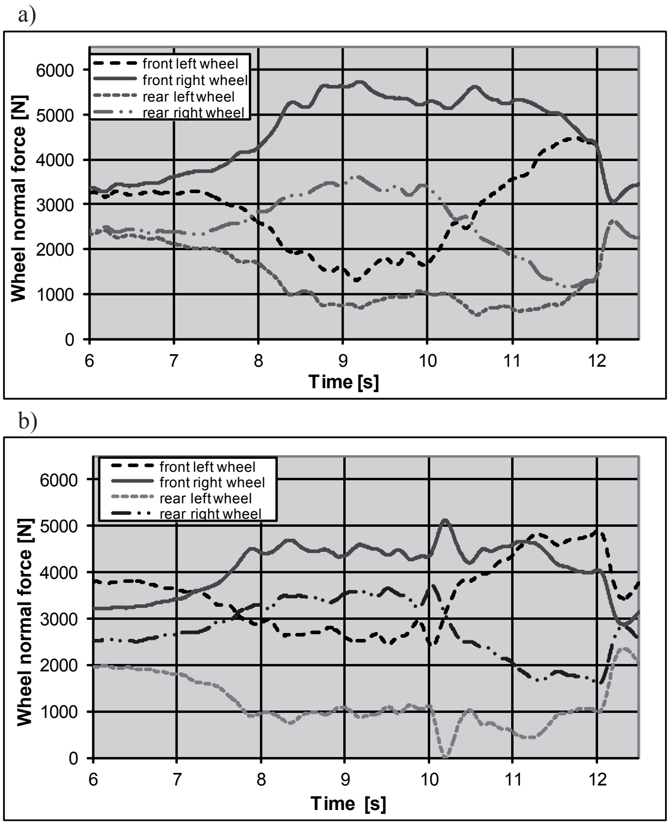


Fig. 11. Wheel load of the vehicle during test to brake on a curvilinear section of the road (the inflation pressure in the right front tire was nominal (a) and reduced (b))

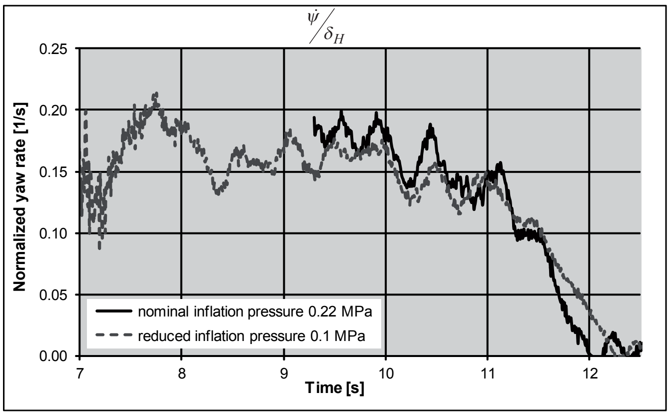


Fig. 12. The normalized yaw rate $\dot{\psi}/\delta_H$ *(right front tire inflation pressure: nominal and reduced)*

The graph (Fig. 11) shows that the decrease of normal force in the front right wheel and changes significantly normal forces affect the behavior of the vehicle during a maneuver. Lowering the front right wheel normal force is compensated by changing normal forces on the rear wheel right and front left. Correction of the left rear wheel load is a bit smaller. It causes the behavior of the vehicle during the test, as shown in the following charts.

Figure 12 shows the effect of lowering the inflation pressure in the tire for the normalized yaw rate. Lowering the standard yaw rate shows a magnification of the vehicle understeer, which confirms the standardized yaw angle of graph shown in Figure 13.

The analysis of the graph that the yaw rate, in the case of executing a vehicle braking maneuver on curved track, will be smaller in the case of motion of the vehicle tire with reduced inflation pressure, and will therefore be at the same angle of rotation of the steering wheel moves along a track with a larger radius .

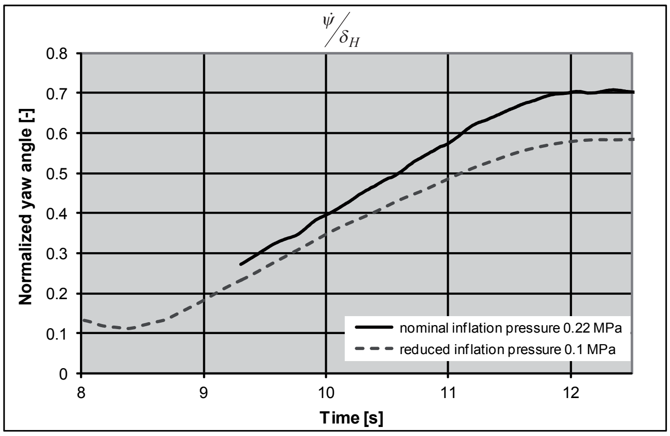


Fig. 13. The normalized yaw angle $\dot{\psi}/\delta_H$ *(right front tire inflation pressure: nominal and reduced)*

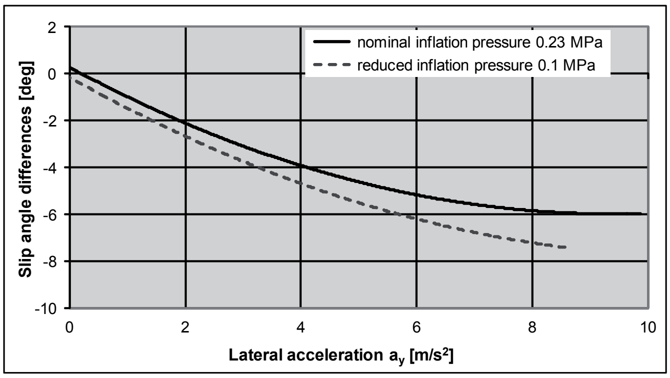


Fig. 14. The slip angle differences on front and rear wheels (right front tire inflation pressure: nominal and reduced)

The vehicle understeer determines difference of slip angles of front and rear wheels. From the graph (Figure 14) clearly shows that the vehicle with reduced inflation pressure in right front wheel has a much greater slip angles difference in the front and rear wheels, and is more understeer throughout the range of acceleration. Determined on this basis, the value of the model parameter MF, $q_{Mz3} = -0.091$ for inflation pressure of 0.1 MPa.

5. Conclusions and discussion

Based on the above analysis and the results of the measurements can be seen that reducing the inflation pressure in the tire results in:

- Change of the vertical replacement stiffness resulting from the stiffness of tires and wheel suspension. Reducing the stiffness is greater at the lower the inflation pressure in the tire. This reduces the lateral stiffness of the tire.
- Increased stopping distance of the vehicle, both on the track straight and curved.
- Increase vehicle understeer, especially during turning maneuvers that increase the normal load of wheel with reduced inflation pressure.
- Change of the tire model parameters MF (magic formula) on the effect of reducing pressure on the tire stiffness characteristics, explicitly for the coefficients associated with radial and lateral stiffness and less important for longitudinal stiffness.

With the tests and analyzes show that the effect of tire inflation pressure has a significant impact on the behavior of the vehicle during braking maneuver and thus the safety of the road traffic. Typically, lowering the inflation pressure in one of the tires is not enough appreciated by road users, shown above tests indicate a significant deterioration in control of the vehicle and increase stopping distances.

References

1. Andrzejewski R. Dynamika pneumatycznego koła jezdnego. Warszawa: WNT, 2010.
2. Arczyński St. Mechanika ruchu samochodu. Warszawa: WNT, 1993.
3. Besselink IJM, Schmeitz A J C, Pacejka, H B. An improved Magic Formula/Swift tyre model that can handle inflation pressure changes. Vehicle System Dynamics, 2010; 48,1: 337–352.
4. Clark SK and others, Machanics of pneumatic tires. Washington DC: National Buremu of Standards, 1971: Monograph 122.
5. Fundowicz P. Droga hamowania na łuku drogi. Warszawa: Politechnika Warszawska, Zeszyty Instytutu Pojazdów; 2010; 1: 103–110.
6. Gillespie TD. Fundamentals of vehicle dynamics. Warrendale: SAE Inc., 1992.
7. Grzegożek W. Modelowanie dynamiki samochodu przy stabilizującym sterowaniu siłami hamowania. Kraków: Zeszyty Naukowe Politechniki Krakowskiej, 2000; Seria Mechanika, monografia 275.
8. Jazar R. Vehicle Dynamics, Theory and applications. Springer Science + Bussines Media, 2009.
9. Lacombe J. Tire model for simulations of vehicle motion on high and low friction road surfaces. San Diego CA: 2000; Proceedings of the 32nd conference on winter simulation.
10. Nijmeijer H, Schmeitz AJC, Besselink IJM. Enhancing the MF-Swift Tyre Model for Inflation Pressure: I.B.A. op het Veld, Eindhoven, 2007;11.
11. Pacejka HB. Tire and vehicle dynamics. Warrendale, SAE, 2006.
12. Parczewski K, Wnęk H. Utilization of the car model to the analysis of the vehicle movement after the curvilinear track. Eksploatacja i Niezawodnosc – Maintenance and Reliability, 2010; 4: 37–46.
13. Parczewski K, Wnęk H. Wpływ niesprawności zawieszeń na stateczność ruchu pojazdu - porównanie badań symulacyjnych i pomiarów. Archiwum Motoryzacji, 2006; 2: 159–169.
14. Parczewski K. Exploration of the shock-absorber damage influence on the steerability and stability of the car motion. Journal of KONES, Powertrain and Transport, 2011; 18-3: 331–338.
15. Pillai PS. Efect of tyre overload and inflation preasure on olling loss (resistance) and fuel consumption of automobile and truck/bus tires. Indian Journal of Engineering & Materaial Science, 2004; 11: 406–412.
16. Prochowski L, Unarski J, Wach W, Wicher J. Podstawy rekonstrukcji wypadków drogowych. Warszawa: WKL, 2008.

Krzysztof PARCZEWSKI, Ph.D. (Eng.)
Department of Internal Combustion Engines and Vehicles
University of Bielsko-Biała
ul. Willowa 2, 43-309 Bielsko-Biała, Poland
e-mail: kparczewski@ath.bielsko.pl

Article citation info:
SKOWRONEK K, WOŹNIAK A. FFT-PCA-LDA classifier in A.C. Generator diagnostics. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 140–146.

Konrad SKOWRONEK
Artur WOŹNIAK

FFT-PCA-LDA CLASSIFIER IN A.C. GENERATOR DIAGNOSTICS

KLASYFIKATOR FFT-PCA-LDA W DIAGNOSTYCE ALTERNATORA*

The methods of A.C. generator diagnostics are discussed. The need of developing new methods is justified. A new classification method is presented that is used for diagnostics of A.C. generator damages. Features of the method are specified. Functioning of the method is analyzed based on examination of damages of A.C. generator diodes. The method is compared with other methods of electric machine diagnostics used in practice.

Keywords: *A.C. generator diagnostics, classification, hybrid method.*

Omówiono metody diagnostyki alternatorów. Uzasadniono konieczność konstrukcji nowych metod. Zaprezentowano nową metodę klasyfikacyjną wykorzystaną do diagnostyki uszkodzeń alternatora. Przedstawiono cechy metody. Działanie metody przeanalizowano na podstawie badania uszkodzeń diod alternatora. Metodę porównano z metodami diagnostyki maszyn elektrycznych stosowanymi w praktyce.

Słowa kluczowe*: diagnostyka alternatora, klasyfikacja, metoda hybrydowa.*

1. Introduction

The problem of A.C. generator diagnostics has a topical meaning, taking into account the needs of automotive industry and motor car users. In the first case the diagnostics is aimed at assessing the product during manufacturing process, in the other – during its operation.

In practice the A.C. generator diagnostics in the vehicle or beyond it is carried out with the following methods [12]:

- comparative,
- oscilloscopic,
- voltmetric,
- with the use of specialized diagnostic devices, for example an indicator device.

In the comparative method the degree of compliance of the measured and standard characteristics of the generator is checked. It is a laborious and inaccurate method. It is rather an individual solution that enables evaluation of the A.C. generator, nevertheless, it does not allow to identify the type and place of the damage. It is equivalent to consideration of the machine as a "black box", without visualization of the connections and without access to the terminals. Statistics of recorded cases does not enable concluding, for example on the type of the manufacturing error.

The most common practical method is a so-called oscilloscopic method. It consists in comparison of standard signal oscillograms with the signal patterns obtained for the considered A.C. generator. This method enables identifying both the locations of manufacturing errors and damages, as for example short-circuits of stator winding to earth. The main signals are usually the output, phase, and interphase voltages. The signals are measured in various points, according to the generator type, in particular in the points where measuring probe access is possible. Therefore, the method itself is not general enough. Moreover, it does not allow to discern single damages and manufacturing errors.

In the voltmetric method the same signals that are estimated in the oscilloscopic method are measured with voltmeter. Therefore, it has the same bad and good points.

The use of specialized diagnostic devices is delimited to a given A.C. generator type. Usually, the access to definite generator terminals should be accessible, particularly to the phase windings ones. This simple method may be easily used even by imperfectly trained workers and, at the same time, it ensures detection of the damages and errors but does not allow to discern them.

As opposed to diagnostics of induction motors, the literature delivers rather few information on more sophisticated diagnostic methods related to A.C. generators. For example in [5] a computer analysis of A.C. generator is proposed, with the use of artificial neural network. Good identification results have been achieved only for some types of errors and damages. Literature delivers no works on:

- proposals of algorithms that allow for assessing the state based on many signals, under dynamic changes of operation conditions;
- proposals of algorithms that perform the classification and regression tasks in the state assessment;
- proposals of algorithms that operate correctly even in extremely different cases;
- adaptation of the model used for the assessment to various tasks: in the frequency and time domains (adaptation of the number of variables and model parameters);
- proposals of a method of forecasting the pattern of selected parameters of the device operation.

Therefore, a proposal of a new method of A.C generator diagnostics, at the assembly belt or located in vehicle, is considered to be purposeful. Such a method should be universal, should be able to

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

detect and classify both the groups of the damages and single damage, allowing, at the same time, to indicate their locations.

2. Analysis of possible solutions

In order to automate and standardize the assessment of the state of new and used A.C. generators the diagnostic results should be free of the components that depend on non-objective factors. At the same time, the method should be sufficiently elastic, which means that it should allow for formulating objective standards of deviations from rated patterns. For example, the damage standard patterns should be possible to formulate. In this sense among the most important methods the ones that make use of random description of the signals are reckoned.

There are many diagnostic methods [2]. Among the most general ones the statistical classification methods should be mentioned, as PCA, ICA, LDA, B&B, SFS, SBS, SFFS&SBFS, SVM, Bayes, LDC, k-NN, and others. In order to analyze large data files recorded during the measurements many methods are used, e.g. the ones used for teaching the patterns [9]. Common distinction between the teaching with or without supervision is made. The first ones includes, among others, PCA, PP, SOM (Kohonen maps) etc. The others are usually classified among the calibration, discrimination, and classification ones, according to the considered problem.

In the discussed case, i.e. in A.C. generator diagnostics, a large number of strongly correlated data must be searched, due to physical structure and the place of operation of these devices. Therefore, preliminary comparative study of selected methods has been carried out. For this purpose the identification of the state of the object or estimation of the values of its operation parameters has been carried out with the use of the following methods: PCA (principal component analysis), LDA (linear discriminant analysis), SVM (support vector method), SNN (artificial neuron networks), MARSPlines (multivariate adaptive regression with the use of splines).

For purposes of searching large data files the PCA method (principal component analysis) appears to be optimal. PCA is a typical teaching method without supervision, based on linear algorithm of feature extraction, with reduction of the data files.

On the other hand, LDA method (linear discriminant analysis) appears to be optimal in case of classification of data groups. The linear discriminant analysis is reckoned among the most frequently used in engineering. It serves for formulating linear discriminant functions. They are created for the samples of a model file. The samples formulate the initially defined groups. The discriminant functions created for them are used for classifying new "virtual" samples for one of the considered groups.

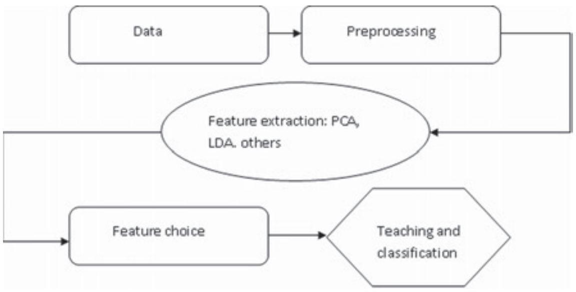


Fig. 1. The stages of statistical classification

Both above mentioned methods (their location with regard to the chain of diagnostic events is shown in Fig. 1) are particularly useful for the data distinguished by Gaussian distribution. Effectiveness of simultaneous use of both these methods for processing of the sample files was noticed, among others, in [11].

One of possible ways of formulating a method distinguished by the above mentioned advantages consists in joining both methods at various classification stages and, at the same time, supplementing it by an algorithm allowing for easier formulating of the standards. This goal may be attained e.g. by wavelet transformation or Karhunen-Loeve transformation. Taking into account common use of FFT algorithms, implementation the Fourier transformation was used in the present paper.

Hence, a new hybrid method FFT-PCA-LDA of damage classification of automotive alternator is proposed, based on the analysis of a reduced data of variables in frequency domain, with the use of multi-dimensional data analysis method. The frequency analyses used in diagnostics are described, for example, in [3, 6, 7, 8, 10]. The paper [7] presents detection of diode damage with the help of a filter tracing appearance of a definite frequency component. Nevertheless, they are distinguished by important constraints. Moreover, as it was mentioned above, large data files must be then analyzed. Therefore, the data should be initially examined with a view to optimize the quantity and contents of the information.

Summarizing, formulation of the new method was carried out with consideration of the diagnostic methods used for typical electric machines, and, at the same time, with the use of the methods applied in diagnostics of the vehicle electric equipment. Such an approach has not been proposed before in the literature related to these problems.

The most frequent damages of A.C. generators and other electric machines are the damages of bearings. Another damage that relatively often occurs in A.C. generator is the damage of the bridge-rectifier. Taking it into account the classification method proposed in the present paper is verified for the bridge-rectifier damages.

2. FFT-PCA-LDA Classifier

In the on-board diagnostic systems the hardware and software solutions are missing, that could clearly determine the generator damage type. The codes delivered by manufacturer or another economical units provide the information on barely several damages (Table 1). Therefore, an attempt is made aimed at a new approach to A.C. generator diagnostics that may be used at measuring stand or for purposes of vehicle on-board diagnostics.

The following assumptions are made:

- the use of signal frequency analysis;
- concluding based on voltage or current signal of the A.C. generator;
- construction of a statistical model using multi-dimensional data analysis for reducing variable dimensions and recognizing the damage standards;
- minimization of the signal acquisition points (non-invasive for the A.C. generator structure.

Table 1. Example codes of A.C. generator damages

Code	Damage description	System
1117	Load signal from DF terminal of A.C. generator	VAG
1209	Rotation signal – A.C. generator terminal	VAG

It should be noticed that various types of the bridge-rectifiers should be discerned only based on output signals: e.g. voltage or current. For comparison – in case of the oscilloscopic method such a discernment is impossible (Fig. 2), since the zero point voltage of generator stator would be necessary for this purpose. Moreover, according to Fig. 2, knowledge of generator structure or, at least, standard course of the signal, would be required.

Similarly, FFT analysis with comparison of the signals does not provide expected results (Fig. 3). According to the measurements

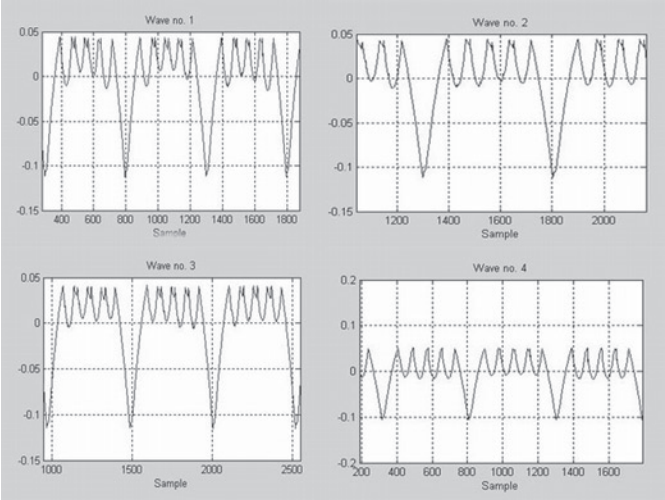


Fig. 2. Voltage time patterns of the damages: A(+), A(−), B(+), and dB(−), the damage of one diode

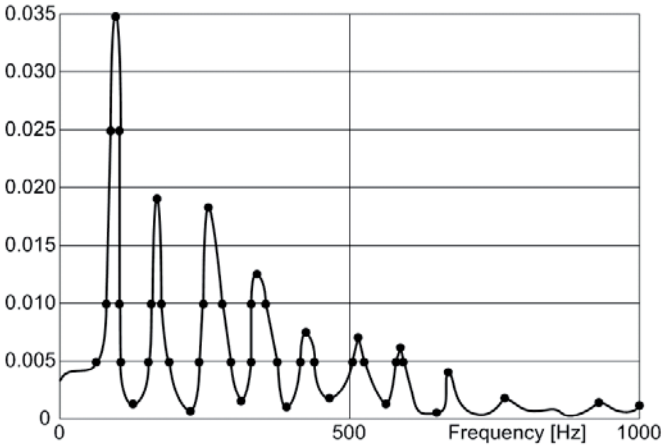


Fig. 3. FFT analysis of the damages: A(+), A(-), B(+), and dB(-), the damage of one diode

(Fig. 2) discernment of the damage types (and even the damage states only) remains impossible in this case. The above problems may be also related to clear randomness of measured patterns of all the signals (Fig. 4).

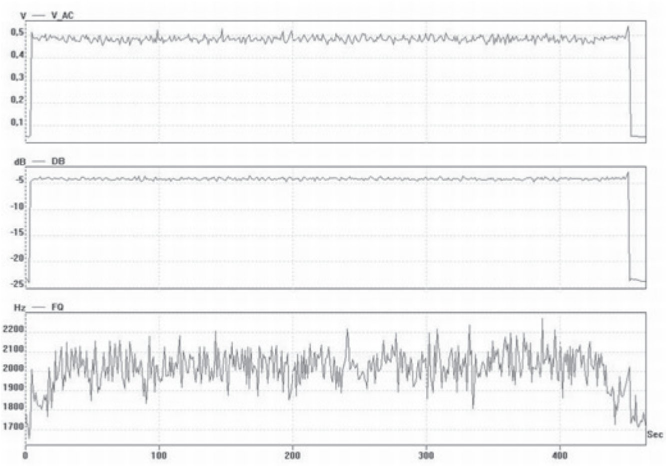


Fig. 4. True A.C. generator signals: from top to bottom: voltage, noise, and frequency

During classification with the method of linear discriminant analysis (LDA) with preliminary extraction of signal features, with the use of principal component analysis (PCA), the vector of variables is automatically chosen. This vector is then considered as a basis for classification of the object to a definite group of standard states. A file of N-samples in n-dimensional space is considered and it is assumed that every image belongs to one of K-classes $\{C_1, C_2, ..., C_K\}$. N_j is the sample number within the class C_j, $u_j = (1/N_j)\sum_{x\in C_j} x$ is the average of the image from the class C_j, $u = (1/N)\sum_{j=1}^{K}\sum_{x\in C_j} x$ is the average of the image from all the samples. The dissipation matrix in the class is given in the form:

$$S_w = (1/N)\sum_{j=1}^{K}\sum_{x\in C_j}(x-u_j)(x-u_j)^T = \Phi_w\Phi_w^T \tag{1}$$

On the other hand, the dissipation matrix between the classes is defined as:

$$S_b = (1/N)\sum_{j=1}^{K} N_j(u_j-u)(u_j-u)^T = \Phi_b\Phi_b^T \tag{2}$$

The dissipation matrix has a form:

$$S_t = (1/N)\sum_{j=1}^{K}\sum_{x\in C_j}(x-u)(x-u)^T = \Phi_t\Phi_t^T = S_w + S_b \tag{3}$$

In case the S_w matrix is not singular the LDA method attempts to determine a projection $W_{opt}=(w_1, w_2,…., w_L)$, that meets the Fisher criterion:

$$W_{opt} = \arg\max_W \frac{\left|W^T S_b W\right|}{\left|W^T S_w W\right|}, \tag{4}$$

where $w_1, w_2, …, w_L$ is the vector of eigenvalues $S_w^{-1}S_b$ with reference to L (≤K−1) the biggest eigenvalues $\lambda_1, \lambda_2,…, \lambda_L$.

In case the S_w matrix is singular its inverse does not exist. The PCA method [6] is then used for projection of the vector of variables on the space of lower dimension in order to remove the singularity.

Principal component analysis (PCA) consists in the use of a statistical algorithm with a view to separate a small number of coefficients that represent the best a large number of properties of extremely large datafile. It is at present very often used, e.g. in MatLab packages, since it enables graphical approximate assessment of datafile structures. In the PCA method the principal components (PC) are generated, that make a part of optimal linear combinations representing the weighted sums of the input data. The combinations result from gradual transforming of the coordinated system with regard to consecutive decreasing values of data variations. These consecutive coordinate axes make a file of further principal components.

The proposed method is performed in two steps: training of the model and concluding. The classifier is called a linear one, taking into account the methods of description of the relationships between the variables, that make use of linear functions (PCA and LDA). The trend of the damages in the analysis carried out based on the covariance matrix is also of linear character. Finally, the PCA ensures maximization of the variation while LDA – separation of the classes. Proper application of LDA method occurs, when [9]:

- distribution of the objects within each sample group approximates the normal distribution;
- the sample groups are linearly separable;

• the variance-covariance matrices of each of the sample groups are comparable;
• total number of the objects must be at least three times as large as the number of the variables.

It was found that the above assumptions correspond to the conditions of data collecting in case of A.C. generator tests.

In order to generate the standards the FFT analysis has been used, with narrowed frequency windows. Analysis of the initial "raw" data consists only in collecting the damage standards with the help of FFT analysis with rectangular window. Preliminary processing is made by multiplying the obtained signal with the Hanning window and, additionally, averaging of the values so obtained for particular frequency components for restricted number of the windows with restricted number of points.

4. Measurements and verification results

The proposed classifier was verified during the tests on the measurement stand shown in Fig. 5. It was provided, among others, with a 4-channel music card ESI Quata Fire 610, measurement converter SENSOR AMP-4ICP, vibration sensor DYTRAN, and the A.C. generator testing stand. A bridge-rectifier damage was simulated for 1 diode and 2 diodes. The damage patterns from the time window for 4000 samples have been collected. The time window was obtained based on FFT analysis for 250 points. The samples were collected for the speeds 800 and 1000 rpm of the A.C. generator shaft.

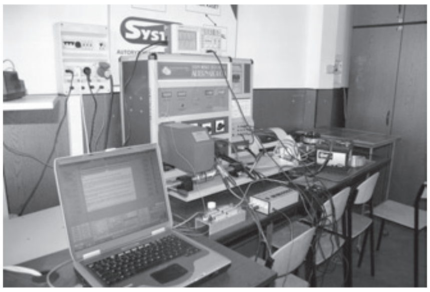


Fig. 5. Measurement stand of the A.C. generators

The damage cases were classified into groups and files. The teaching standards were formulated for the damages in particular stator phases, for the damages of positive and negatives diodes, and the damages of particular phases and polarity. It was found during the tests that eigenvalues of the covariance matrix for each of examined cases approximated each to other and were distributed similarly as in the case presented in Fig. 6.

It was found during the tests of discernment of one diode damage at the level of polarity and with varying speed, that the results are good. Nevertheless, correctness of classification significantly depends

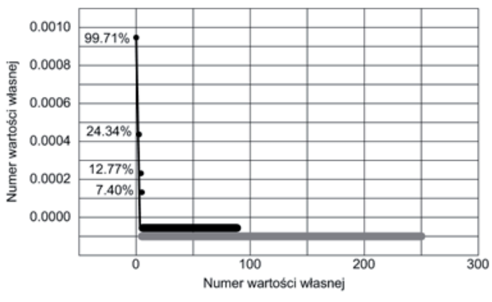


Fig. 6. Eigenvalues of the covariance matrix for one damaged diode at the speed equal to 800 rpm

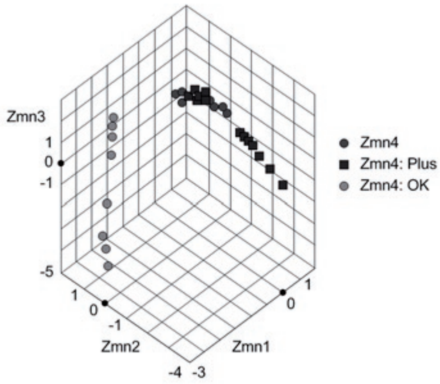


Fig. 7. Categorized 3D plot of principal components for one damaged diode at the speed equal to 800 rpm

Table 2. Percent quality of classification for one damaged diode at the speed equal to 800 rpm

State	Percent
OK.	100,00
Fault_Minus	66,67
Fault_Plus	90,00
Total	81,43

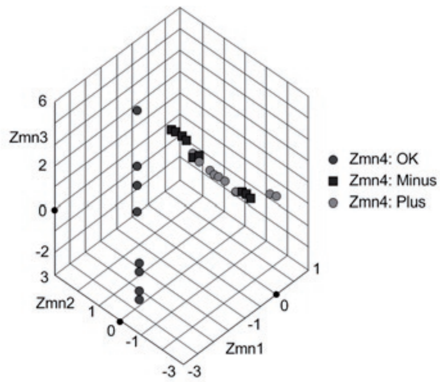


Fig. 8. Categorized 3D plot of principal components for one damaged diode at the speed equal to 1000 rpm

Table 3. Percent quality of classification for one damaged diode at the speed equal to 1000 rpm

State	Percent
OK	100,00
Fault_Minus	66,67
Fault_Plus	66,67
Total	71,43

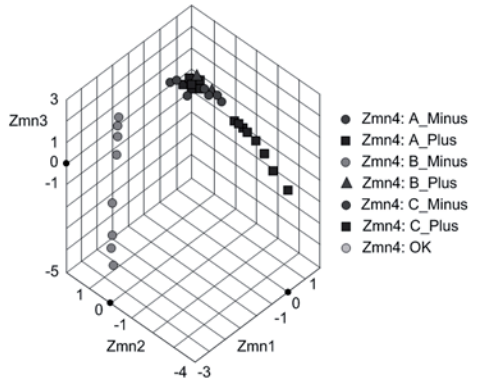


Fig. 9. Categorized 3D plot of principal components for one diode damaged at polarity and phases level at the speed equal to 800 rpm

Table 4. Percent quality of classification for one diode damaged at polarity and phases level at the speed equal to 800 rpm

State	Percent
OK	100,00
Fault_AMinus	100,00
Fault_APlus	70,00
Fault_BMinus	100,00
Fault_BPlus	90,00
Fault_CMinus	100,00
Fault_CPlus	100,00
Total	94,28

on the speed (Figs 7 and 8). On the other hand, in all tested cases the trend for the variables processed with FFT-PCA-LDA was clear.

Taking into account unsatisfactory classification results (Tables 2 and 3) with the use of polarity or phase only, the damages were classified with simultaneous consideration of both these signals. In consequence, it appeared that effectiveness of classification grew significantly. For example, Fig. 9 and Table 4 present the results of discernment tests of a damage of one diode at polarity and phases level at the speed equal to 800 rpm.

Interesting results have been obtained in examination of the damages of more than one element in the same phase. In all these cases 100-percent effectiveness was achieved, that is shown, for example, in Fig. 10 and Table 5.

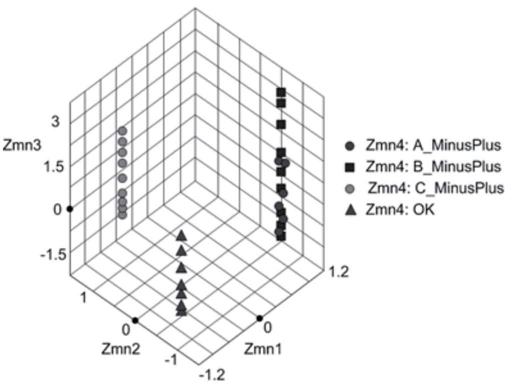


Fig. 10. Categorized 3D plot of principal components for 2 diodes damaged for the same phase at the speed equal to 800 rpm

Table 5. Percent quality of classification for 2 diodes damaged for the same phase at the speed equal to 800 rpm

State	Percent
OK	100,00
Fault_A	100,00
Fault_B	100,00
Fault_C	100,00
Total	100,00

It was found that the best results are obtained with the use of the FFT-PCA-LDA classifier for the signals referring to polarity and phase for more than one damage.

5. Classification results obtained with other methods

5.1. Artificial neural networks

In order to confirm good quality of the proposed classifier similar tests have been carried out for other selected classification methods.

Comparison of the classification process was performed for the following network types: linear, PNN, RBN, three-layer perceptron, and four-layer perceptron. The tests have been made for 20 of these network types and for the cases of the files including non-reduced and reduced variables. For the non-reduced file – 250 variables – with one diode damaged at the polarity and phases level, the best result was achieved for linear network 248:248-7:1, with teaching quality 1.000000; validation quality 0.705882; testing quality 0.823529; teaching error 0.000000; validation error 2.254002; testing error 2.254002. Diagram of the winning network is shown in Figure 11.

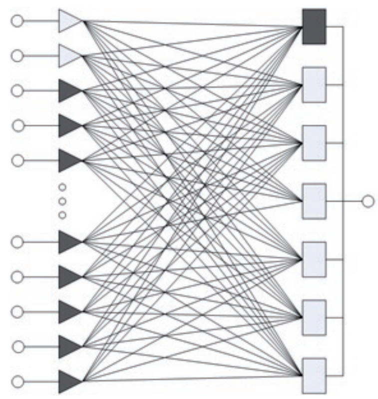


Fig. 11. Diagram of linear network

For the reduced file – 3 variables (principal components) – with one diode damaged at the polarity and phases level, the best result was achieved for the network MLP 3:3-9-9-7:1, with teaching quality 1.000000; validation quality 0.941176; testing quality 1.000000; teaching error 0.042153; validation error 2.763365; testing error 0.048920. Diagram of the winning network is shown in Figure 12

For the reduced file – 3 variables (principal components) – with one diode damaged at the phase level, the best result was achieved for the network RBF 3:3-7-4:1, with teaching quality 0.888889; validation quality 0.823529; testing quality 0.823529; teaching error 0.232214; validation error 2.030926; testing error 5.351756. Diagram of the winning network is shown in Figure 13.

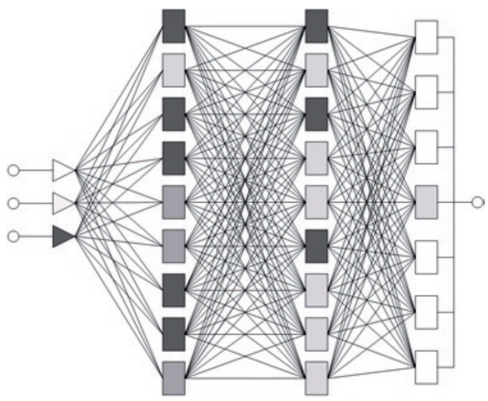


Fig. 12. Diagram of MLP network

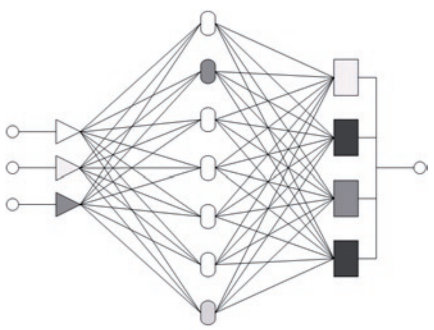


Fig. 13. Diagram of RBF network

Table 6. Results of assignment

Concentration number	Number of the cases
1	20
2	3
3	10
4	2
5	20
6	8
7	7

5.2. Grouping with k-average method

Classification for a reduced datafile – 3 variables (principal components) was carried out for 70 cases. The cases were interrelated with k-average method, missing data were complemented with the cases. 7 concentrations were separated, while the solution was found after 2 iterations. The results are shown in Table 6.

6. Summary

Automotive A.C. generator is an electric machine operating in specific conditions that vary cyclically during the operation, similarly to the device load. The methods that enable accurate identification of the generator state may be applied only after its disassembly from the vehicle.

Preliminary results of the tests and measurement analysis with the use of FFT analysis of the output voltage signal of the generator and with data mining show that the hierarchical FFT-PCA-LDA classifier is very effective in discernment of the damages of the generator bridge-rectifier. These damages are distinguished by linear trend in 3D space of principal components for the damage samples.

The data mining methods allow for using "former" data, many times analyzed before. At present most of the diagnostic systems, aimed not only at vehicle diagnostics, are based on the use of artificial neural networks, that are very successful in the regression or classification problems. The proposed hybrid FFT-PCA-LDA classifier, that is characterized by good discretization ability of the damages distinguished by a linear trend, with remarkably more simple adjusting process as compared to other classifiers. The classifier enables reducing the number of the variables down to 3, maintaining accurate operation. The highest number of the data erroneously classified by this classifier occurred for the damages converted to the groups (e.g. the damages in particular phases). In case of all the damages the data have been correctly assigned to particular classes. The results obtained enable further analysis. The classifier should be checked for various values of rotational speed and load, since increased number of the signals improved the classification quality. Further tests will be aimed at the use of correlation of mechanical and electrical signals of the A.C. generator. The number and type of the signals will depend, among others, on the fact whether the A.C. generator is to be tested at the assembly belt or in vehicle. In the first case detection of possibly large number of faults is important, while in the other – stating its fitness for use.

References

1. Alternator test apparatus and method, United States Patent, 4 178 546, 1979.
2. Chen VC. Evaluation of Bayes, ICA, PCA and SVM Methods for Classification, RTO-MP-SET-080, 2004.
3. Dziubiński M. Koncepcja diagnostyki wyposażenia elektrycznego samochodu. Eksploatacja i Niezawodność – Maintennance and Reliability 2002; 2 (43): 39–47.
4. Haardle W, Simar L. Applied Multivariate Statistics Analysis, MDTech, 2003
5. Jastriebow A., Gad S, Słoń G. Analiza komputerowa diagnozowania defektów alternatora z wykorzystaniem sztucznej sieci neuronowej. Diagnostyka 2002; 27: 7–10.
6. Jolliffe IT. Principal Component Analysis, Springer Verlag, 2002.
7. Kościelny JM. Diagnostyka zautomatyzowanych procesów przemysłowych, Exit, Warszawa 2001.
8. Rawicki S, Stein Z, Stolpe M, Zielińska M. Analiza harmonicznych w prądach i napięciach samochodowej prądnicy prądu przemiennego (alternatorze) z układem prostownikowym. ZKwE 1999: 567–570.
9. Stanimirova M, Daszykowski B, Walczak. Metody uczenia z nadzorem – kalibracja, dyskryminacja i klasyfikacja, w D. Zuba, A. Parczewski,

Chemometria w analityce, IES, Kraków, 2008.
10. Stein Z, Zielińska M. Problems of electromagnetic compatibility of Automotive alternating current generator (alternator), AMTEE, 2005, s. H13–H17.
11. Sudarshan S. Deshmukh, Pooja S. Deshmukh, Sachin S. Sawalkar, Pose Variant Based Comparative Analysis of PCA and LDA, Proceeding ICETET '09 Proceedings of the 2009 Second International Conference on Emerging Trends in Engineering & Technology IEEE Computer Society Washington, DC, 2009:188–191.
12. Źółtowski B, Tylicki H. Osprzęt elektryczny pojazdów mechanicznych, ATR Bydgoszcz, 1999.

Prof. Konrad SKOWRONEK, Ph.D. hab. (Eng.)
Faculty of Electrical Engineering
Poznań University of Technology
ul. Piotrow 3a, 60-965 Poznań, Poland
e-mail: Konrad.Skowronek@put.poznan.pl

Artur WOŹNIAK, M.Sc.
Wood-Mizer Industries Sp. z o.o.
ul. Nagórna 114
62-600 Koło, Poland
e-mail: artur_wozniak@o2.pl

Article citation info:
YANG Z-J, CHEN C-H, CHEN F, HAO Q-B, XU B-B. Reliability analysis of machining center based on the field data. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 147–155.

Zhao-jun YANG
Chuan-hai CHEN
Fei CHEN
Qing-bo HAO
Bin-bin XU

RELIABILITY ANALYSIS OF MACHINING CENTER BASED ON THE FIELD DATA

ANALIZA NIEZAWODNOŚCIOWA CENTRUM OBRÓBKOWEGO W OPARCIU O DANE TERENOWE

Machining center is the complex machinery, with high level automation and complicated structures, so there are lots of failures. When a random failure occurs, the failed machining center stops and causes a production line or even the whole workshop to stop functioning. The frequent failure leads to the low levels of reliability and production rate. In order to help users and manufacturers optimize maintenance policy to improve the reliability for machining center, this paper presents descriptive statistics of the failure data and develops the failure trend using power-law process, simultaneously establishes the routine inspection and regular inspection as well as the sequential preventive maintenance under maintenance cost constraints. The proposed model could be a useful tool to assess the current conditions, predict reliability and optimize the machining center maintenance policy.

Keywords: *failure analysis, machining center, maintenance policy, power-law process, repairable system.*

Centrum obróbkowe to skomplikowany mechanizm o wysokim poziomie automatyzacji oraz złożonej konstrukcji, w związku z czym ulega licznym uszkodzeniom. Przy wystąpieniu przypadkowej awarii, uszkodzone centrum obróbkowe przestaje działać i powoduje zatrzymanie linii produkcyjnej a nawet całego oddziału produkcyjnego. Częste awarie obniżają poziom niezawodności oraz tempo produkcji. Aby pomóc użytkownikom i producentom zoptymalizować politykę utrzymania ruchu w celu poprawy niezawodności centrów obróbkowych, w niniejszym artykule przedstawiono statystyki opisowe dotyczące danych o uszkodzeniach i opracowano trend uszkodzeń w oparciu o proces spełniający prawo potęgowe. Jednocześnie ustalono zasady rutynowej inspekcji i okresowych przeglądów, jak również sekwencyjnej obsługi zapobiegawczej przy ograniczonych wydatkach na utrzymanie ruchu. Proponowany model może być użytecznym narzędziem dla potrzeb oceny aktualnych warunków oraz przewidywania niezawodności w celu optymalizacji polityki utrzymania ruchu centrum obróbkowego.

Słowa kluczowe*: analiza uszkodzeń, centrum obróbkowe, polityka utrzymania ruchu, proces spełniający prawo potęgowe, system naprawialny.*

1. Introduction

With the increasing development of high-speed and high-precision technologies, machining center is becoming the main equipment for advanced manufacturing technology. It is a typical electromechanical product mainly composed of mechanics, electronics and hydraulics, etc. In most cases, machining center is usually used in production lines for mass production, thus it fails more often than NC lathe [14].

Machining center is often regarded as a repairable system, so it can be restored to an operational state by some maintenance actions such as corrective maintenance (CM) and preventive maintenance (PM) [5, 6]. There is a failure-repair-failure cycle with the ability to repair a failed repairable system. Depending on the features of the repairable system, the distribution of the times to the first failures may not be the same as that of the times between successive failures. Therefore, the traditional life distribution models are not appropriate for the reliability analysis for the repairable system [1, 7].

Failure point process models are characterized by isolated events occurring at instants distributed randomly over a time continuum. So we can use point process models to describe failure process for the repairable system.

Machining center during the whole life acts with many failures which may result in the production of an entire workshop being halted. How does one improve the design of machining center? How does one find out the failure causes for the machining center? Which trends do the failure times follow? Is there an optimal maintenance policy for machining center? The above problems need to be solved to improve the reliability of machining center. This paper studies the failure analysis as well as failure trend model of machining center.

2. Brief description of machining center

The machining center discussed in this paper employs Mitsubishi 64m digital control system with digital AC servo system which has high-precision mode G61.1 and high-speed machining mode G05P3. The CNC system and some electronic components, such as relays, transformer and contactor switches are fixed in the cabinet. The spindle is driven by AC spindle motor, with speed varying from 60 to

8000rpm. The three feeding motions are driven by AC servomotors through ball screws and controlled by CNC simultaneously. In order to raise productivity, there is an automatic tool changer which includes 20 cutters [4].

The machining center is not only appropriate for cutting components such as plate, shaft and rod parts, but also for processing mold parts.

3. Data collection and analysis

3.1. Data collection

Data collection is the basis of failure analysis. The more detailed and truly the failure data is, the more accurate the analysis result is. Tables of operation records and maintenance reports are made in order to collect failure data in a unified format [13]. The operation records table should contain the following information:

1. Product name, product model, product size and manufacturing number.
2. Production date, start date of utilization and valuation date.
3. Other information about operation.

The above information should be recorded in the Fig.1 operation records table.

Product Number:________________ Manufacturer:____________

Production Date:________________ Utilization Date:___________

Valuation Date__________________

Examination Time: ___________to____________

Date	Shifts	Normal/Failure	Operation status		Operator (Signature)	Comment
			Down			
			Down time	Time of restoring		
	Day shift		:	:		
	Middle shift		:	:		
	Night shift		:	:		
			:	:		
			:	:		
			:	:		

Fig. 1. Operation Records Table

The maintenance reports table should contain the following information

1. Failure date and time
2. Failure phenomenon
3. Description of the failure cause
4. Repair process and repair time
5. Other information about machining center failure

The above information should be recorded in Fig.2 maintenance reports table.

The failure data is stored in Excel sheets, and then the time between failures can be obtained by the function of "TEXT (value, format_text)".

Every failure is categorized as spindle system (SS), CNC system (CNCS), electrical system (ES), hydraulic system (HS), tool magazine (TM), lubrication system (LS), screw and guide system (SAGS), servo system (Servo), changeable table (CT), pneumatic system (PS), guard system (GS), cooling system (CS), swarf conveyor (SC) or clamping accessory (CA) based on the function sharing, function independence and convention division principles.

Group__________ Workshop__________ Date(mm-dd-yy):_______

Product Number:		Product Model:
Product Name:		
Failure Time	From: Date: (dd/mm) Time: (hh:mm) To: Date: (dd/mm) Time: (hh:mm)	
Failure Phenomenon		
Failure Position		
Failure Cause		
Repair Process		
Repair Time	Hour/Machine:	Man-Hour:

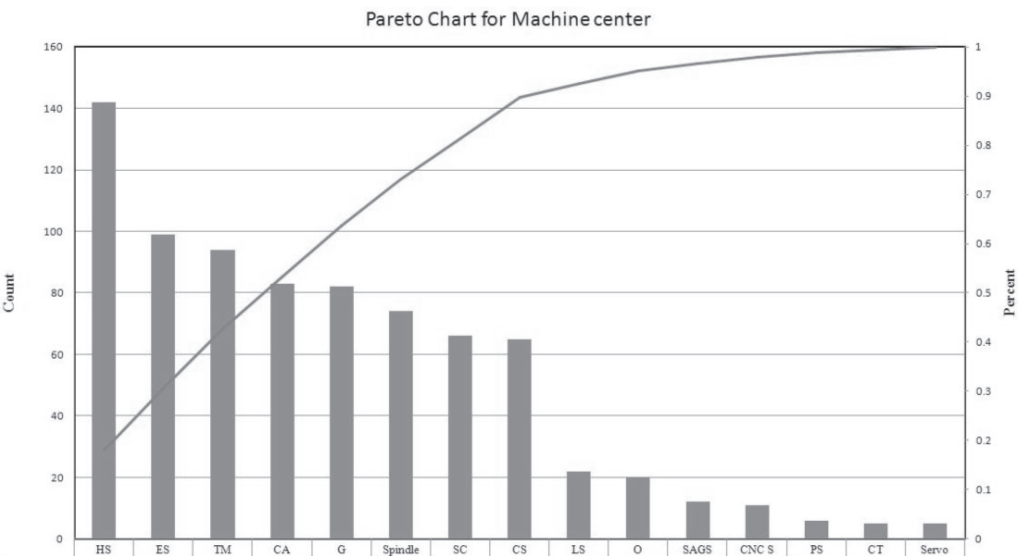
Operator: (Signature):

Fig. 2. Maintenance Reports Table

3.2. Failure analysis

The failure data analyzed in this paper were derived from practical application of twelve machining centers which were manufactured by Dalian Machine Tools Group located in northeast of China. These machining centers were used in a typical representative company of FAW (Fist Auto Works of China) and were traced over the time from 2005 to 2010.

In order to find the weak subsystems, the failures analysis are done and shown in Table 1 which consists of the basic features of repair time, standard deviation (SD) and coefficient of variation (CV). The pareto diagram of the failures is drawn in Fig.3 based on the failure data. In Fig.3, we observe that the HS had the most failures followed by ES, TM, CA, GS and SS and the sum failures of the first six



	HS	ES	TM	CA	G	Spindle	SC	CS	LS	O	SAGS	CNC S	PS	CT	Servo
Count	142	99	94	83	82	74	66	65	22	20	12	11	6	5	5
Cum	0.180661	0.306615	0.426208	0.531806	0.636132	0.730279	0.814249	0.896946	0.924936	0.950381	0.965648	0.979643	0.987277	0.993638	1
Percent	0.180661	0.125954	0.119592	0.105598	0.104325	0.094147	0.083969	0.082697	0.027989	0.025445	0.015267	0.013994	0.007633	0.006361	0.006361

Fig. 3. Pareto diagram of the failures position of machining center

Table 1. Results of the repair time analysis

Subsystem	Repair time	SD	CV	Subsystem	Repair time	SD	CV
HS	1.73	2.54	1.47	CS	2.29	7.68	3.35
ES	0.93	0.68	0.73	LS	0.92	0.28	0.31
TM	1.66	3.35	2.03	SAGS	4.04	13.44	3.33
CA	1.27	1.33	1.05	CNCS	1.37	1.75	0.5
G	1.28	1.29	1.01	PS	0.86	0.14	0.17
SS	2.65	5.17	1.95	CT	0.8	0.45	0.56
SC	0.98	0.80	0.82	SS	1.41	0.28	0.20

subsystems accounted for 73%. Furthermore, 18% of all failures were observed at the HS. From what has been mentioned above, the HS was a large hindrance to the improvement of the reliability. Effectively, with the development of direct-drive technique of the spindle, it had been simplified greatly and the reliability had been raised remarkably. Compared with it, the ES and TM had been improved little.

The CNCS, PS, CT and Servo had few failures seen from Fig.3. Generally speaking, the reliability of these subsystems was much higher than that of the HS, ES.

3.3.1 Failure analysis of HS

The failures of HS accounted for 18% of all the failures, more than any other subsystem. The failures of HS consisted of damages of pumps, solenoids, valves and hoses. The main failure phenomenon and causes are listed in Table 2. From Table 1 and Table 2, the following observations can be made:

a) The main failure causes of the HS were damages of oil pipes, solenoid valves and oil seal indicating that the outsourcing components were unqualified.
b) The HS-related failures required an average repair time of 1.73*h* which was the fourth longest of all the subsystems. The impact on the availability of machining center was significant.

Table 2. Failure phenomenon and causes of HS

Order	Failure phenomenon	Failure causes
1	Oil leaks from oil pipes	Oil pipes damage
2	Tool doesn't changes	Solenoid valves damage
3	Clamping accessory doesn't work	Damage of oil pipes
4	Oil leaks from cylinder	Oil seal wears out
5	Oil pressure is low	Gear pump damages
6	Oil pressure is not stable	The oil goes bad
7	Changeable table turns slowly	The filter is blocked
8	The hydraulic system alarms	The oil temperature is high
9	The oil leaks from pipe joints	The coupling cutting ferrule is loose
10	The position of the cutter holder is wrong	The reversing valves wears out

3.3.2. Failure analysis of ES

12.6% of all the failures were classified as ES. This category included failures of the electrical devices such as switch, lamp-stand, MCB, power module, relays and limit switch. From Table 1 and Table 3 the following observations can be made:

a) The repair time of the ES was the fourth least of all the subsystems.
b) The repair time was 0.93*h*, with low variability because CV was less than one.

Table 3. Failure phenomenon and causes of ES

Order	Failure phenomenon	Failure causes
1	The spindle moves	The position switch damages
2	Work lights doesn't work	Lamp-stand damages
3	The MCB is off	The MCB damages
4	Clamping accessory doesn't work	The switch damages
5	The machining center doesn't work	The power module damages
6	The fan doesn't rotates	The button damages
7	The fuse wire damages	Water goes in fuse box
8	The power line breaks	The power line exposes
9	1020 alarms	A coil of cooler breaks
10	The processing size is out of tolerance	Reset the limit value
11	1027, 1008 and 1007 alarm	Reset he heat transfer element

Table 4. Failure phenomenon and causes of TM

Order	Failure phenomenon	Failure causes
1	The tool can't be loosen	Damage of button
2	Drop of the tool	The tool arm goes down
3	The manipulator can't work	Damage of proximity switch
4	The tool magazine can't rotate	Damage of rolling bearing
5	The tool changes improperly	The wing piece is locked
6	Drop of the tool when machining	The groove of the catch tool wears out
7	The position of the tool sheath is wrong	Damage of spring of location
8	1024 alarms	The time of changing tool is too long

3.3.3. Failure analysis of TM

Failures of TM accounted for 12.9%. Failures of TM consisted of the wrong position of the tool arm, damages of proximity switch and button. TM failures required an average repair time with 1.66*h*. The moderate repair time was mainly due to the long diagnostic time.

Table 5. Failure phenomenon and causes of CA

Order	Failure phenomenon	Failure causes
1	Can't find the centering	Damage of pin
2	The CA can't work	Damage of screw
3	The pressure of clamping is too small	Position of location of clamping is high
4	The CA can't work	Damage of the clamping box
5	Oil leaks from CA	Loose of screw
6	The pressure of clamping is too small	Too much iron chipping

3.3.4. Failure analysis of CA

CA failures accounted for 10.55% of all failures. This category of failures included damages of pin, screw and box of clamping. These types of failures needed the machining center to be shut down resulting in long repair time. From Table 5, we can see that:

a) There were 6 types of failure phenomenon and the main failure causes were damages of mechanical components.
b) The failures of CA mainly caused by the mistakes of the users.

3.3.5. Failure analysis of GS

Table 6. Failure phenomenon and causes of GS

Order	Failure phenomenon	Failure causes
1	The guard breaks away	Deformation of guard seriously
2	The guard breaks away	Too much iron chipping
3	The water sinks in the guard	The bolt of protective door is loose
4	The guard of screw breaks away	The welding of guard breaks
5	Drop of protective door	The bolt of protective door is loose
6	Damage of slide guard	The guard is pulled bad
7	The door of guard fails to open	Damage of the guard roller
8	Poor position of work station	Damage of the guard

GS failures accounted for 10.4% of all the failures. This category consisted of loose of bolt and damage of protective guard. The failure causes were the improper length of guard. The repair time of GS was close to that of CA. It was because the failures of GS and CA were easy to diagnose and repair.

3.3.6. Failure analysis of SS

9.4% of all the failures were spindle failures. Spindle failures had the second longest repair time which was 2.65*h*. The SS was one of the most important subsystems of the machining center. The impact of SS failures on processing parts leaded to the poor precision. The failure phenomenon and causes are listed in Table 7. It is found that the failures mainly caused by the poor assemblage.

3.3.7. Failure analysis of other subsystems

27% of all the failures were failures of other subsystems. These subsystems had fewer failures than the first six subsystems. The reason was the technologies of these subsystems were stable and well understood.

Table 7. Failure phenomenon and causes of SS

Order	Failure phenomenon	Failure causes
1	The processing part has poor precision	Radial endplay of spindle
2	No motion	Damage of motor
3	The parallel of processing part is poor	The bearing clearance is big
4	Poor precision of spindle	Wear out of the bearing
5	The speed of the motor is so low	Lubricants of the spindle is poor
6	The PLC alarms	Damage of spindle disc claw
7	Drop of the cutting tool	Damage of spring
8	The spindle doesn’t rotate	Oiliness of spindle box
9	Abnormal sound in spindle box	Damage of oil cooler
10	700 or 705 alarms	Loose of cable
11	The spindle doesn’t work	Damage of the belt

4. Reliability analysis of machining center

4.1. Model of PLP

Machining center as a repairable system is often modeled by counting processes. A common procedure for analyzing a set of data derived from repairable systems is referred to [8, 10]. The system is observed from instant t=0, and let T_1, T_2 ... T_i denote the successive failure times, X_1, X_2... X_i denote times between failures, thus $X_i=T_i-T_{i-1}$.

The model of NHPP is commonly used in the reliability analysis of complex repairable system with failure intensity function Eq. (1) and cumulative intensity function Eq. (2)

$$\text{h}\ h(t) = \lambda\beta t^{\beta-1} \tag{1}$$

$$H(t) = \lambda t^{\beta} \tag{2}$$

The above intensity function is called the power-law process (PLP). Under the PLP, when β<1, there is positive reliability growth. That is, the system reliability is improving due to corrective actions. When β>1, there is negative reliability growth.

4.2. Analysis of failure data for machining center

4.2.1. PLP of HS

The failure data of machining centers analyzed in this paper were collected from 2005 to 2010. All these machining centers were used in two automotive production lines. So assume that the machining centers had the similar using conditions. Table 8 lists the failure data of HS. Denote variable T_k the k^{th} failure time and t_k is its realization.

Let $0 < t_{i1} < t_{i2} < \ldots < t_{in_i}$ denote the sequential failures times, then the likelihood function under the minimal repair assumption can be shown as [2]

$$L = \prod_{i=1}^{k} \left\{ e^{-\lambda T_i^{\beta}} \prod_{j=1}^{n_i} (\lambda\beta t_{ij}^{\beta-1}) \right\} \tag{3}$$

Where k is the number of machining centers, n_i is the number failures of the i^{th} machining center and T_i is the time-terminated data.

Then the maximum likelihood estimates (MLE) of λ and β are given by

$$\hat{\lambda} = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} T_i^{\beta}}, \quad \frac{1}{\hat{\beta}} = \frac{\sum_{i=1}^{k} T_i^{\hat{\beta}} \ln T_i}{\sum_{i=1}^{k} T_i^{\hat{\beta}}} - \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \ln t_{ij} .$$

In general, these equations cannot be solved explicitly for $\hat{\lambda}$ and $\hat{\beta}$, but can be solved by iterative procedures. Once the estimates $\hat{\lambda}$ and $\hat{\beta}$ are got, the MLE of the intensity function is given by

$$\hat{h}(t) = \hat{\lambda}\hat{\beta} t^{\hat{\beta}-1} \tag{4}$$

Table 8. Failure data of HS

1	2	3	4	5	6	7	8	9	10	11	12
162	1813	2540	19117	8121	17695	10837	8120	7242	3089	1691	370
19627	18296	11831	22629	12439	20530	13785	10328	14535	3393	1702	1039
28576	27232	12449	32312	17707	20743	22018	20384	15705	13508	9800	1472
28583	44728	14830	32528	21390	26365	23529	32960	23551	16981	10233	2345
31710		15204	35197	23776	26920	23559	34501	25741	26278	14000	3419
31811		15496	35432	24756	28205	25740	35527	26673	44728	15130	5190
36227		16428	37278	27702	29453	26605	43798	35038		22047	5239
36246		20502	38214	27727	29753	35037	44728	44728		22854	6369
36827		21323	38411	29758	29993	44728				25740	13286
37016		26278	38910	29824	36538					25740	14093
37037		44728	44143	29846	40900					26983	15101
37086			44582	30436	41825					27871	15132
38100			44728	35169	43357					29621	16132
38125				41481	43359					30016	16978
38140				44728	44728					35035	17379
38270										44728	17859
38388											18221
38435											19109
38482											19732
39110											20859
39299											21254
40286											26273
40616											44728
42433											
43466											
44209											
44728											

Using the Excel solver [11], the PLP of HS is obtained as below

$$h(t) = 2.66 \times 10^{-5} * t^{-0.27} \tag{5}$$

$$H(t) = 1.68 \times 10^{-5} * t^{1.27} \tag{6}$$

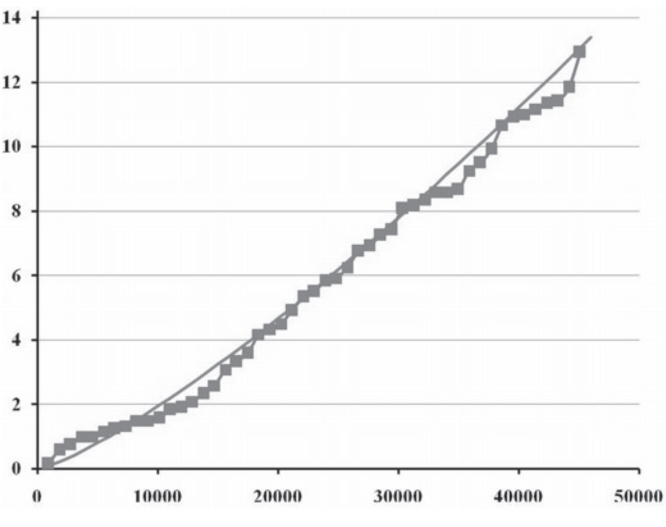


Fig. 4. The cumulative intensity function of PLP for HS data

So the cumulative intensity function is as shown in Fig.4.

4.2.2. Goodness-of-fit test

To determine whether the NHPP is a more appropriate model than the homogeneous Poisson process, a trend test on the failure times is performed [3].

The hypotheses tested are

H_0: The intensity function is constant (β=1).

H_1: The intensity function is not constant ($\beta \neq 1$).

The test statistic is computed from

$$\chi^2 = \frac{2n}{\hat{\beta}} \tag{7}$$

where n is the number of failures and β is MLE the growth or deterioration rate.

Therefore, χ^2 =288/1.27=226.77. Since $\chi^2 > \chi^2_{crit,0.05}$, a significant trend is present. According to the above analysis, to perform the goodness-of-fit for the PLP intensity function, the hypotheses are:

H_0: A PLP with intensity $h(t) = \lambda \beta t^{\beta - 1}$ describes the failure data.

H_1: The above process does not describe the data.

The test statistic is computed from

$$\tilde{\beta} = \frac{n-1}{n}\beta \quad (8)$$

where n is the number of failures.

The Cramer-von Mises [8, 9, and 15] goodness-of-fit test statistic is computed by

$$C_M = \frac{1}{12M} + \sum_{i=1}^{M}[(\frac{t_i}{t_k})^{\beta} - \frac{2i-1}{2M}]^2 \quad (9)$$

where M=n for time-terminated data, t_k=T and T is the total cumulative test time for time-terminated data.

For the data provided in Table 8, the test statistic is computed as follows

$$M=144,\ \tilde{\beta}=1.23,\ C_M = \frac{1}{12M} + \sum_{i=1}^{M}[(\frac{t_i}{t_k})^{\beta} - \frac{2i-1}{2M}]^2 = 0.09$$

The significance α=0.05, the critical value is 0.22. Since C_M<0.22, H_0 is accepted.

4.2.3. PLP of the other subsystems

The machining center consists of thirteen subsystems in series with automated control system. The PLP of the other subsystems can be obtained and listed in Table 9 by the method mentioned in section 4.2.1 and 4.2.2.

Table 9. PLP of subsystems for machining center data

Systems	λ	β	C_M	Cramer-von Mises
HS	1.68×10^{-5}	1.27	0.09	0.22
ES	3.43×10^{-5}	1.17	0.19	0.22
TM	4.08×10^{-3}	0.72	0.12	0.22
CA	2.35×10^{-5}	1.19	0.15	0.22
G	6.40×10^{-5}	1.11	0.21	0.22
Spindle	1.22×10^{-4}	1.02	0.19	0.22
SC	7.65×10^{-6}	1.27	0.21	0.22
CS	1.71×10^{-4}	1.00	0.12	0.22
LS	7.55×10^{-6}	1.24	0.06	0.22
O	1.79×10^{-4}	0.89	0.18	0.217
SAGS	1.73×10^{-4}	0.90	0.17	0.212
CNC S	9.09×10^{-9}	1.77	0.08	0.212
PS	2.01×10^{-7}	1.43	0.09	0.212
CT	1.53×10^{-8}	1.63	0.07	0.199
Servo	1.63×10^{-10}	2.00	0.07	0.199

The critical value of the goodness-o-fit test at 5% significance level is 0.22. Seen from Table 9, almost all statistics C_M are less than 0.22. Therefore, the hypothesis that the models listed in Table 9 can be used to estimate the trends of subsystems, respectively.

4.2.4. Analysis of failure data of machining center

In this section we establish the PLP for the machining center. The failure data were more, so they were not listed here. And the related characteristics are shown in Table 10. Fig.5 shows the cumulative intensity function of machining center.

Table 10. PLP of machining center

λ	β	C_M	Cramer-von Mises
3.41×10^{-5}	1.30	0.17	0.22

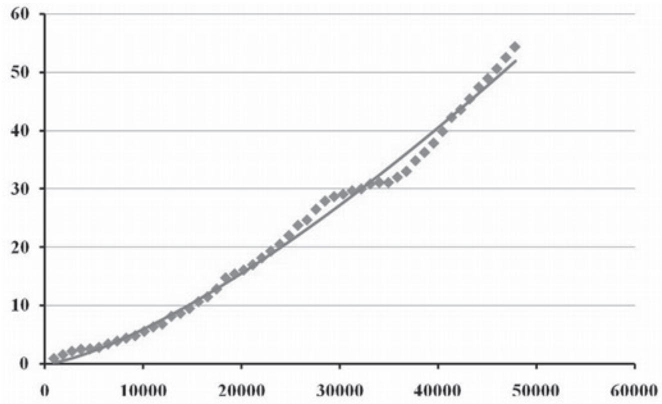


Fig. 5. The cumulative intensity function of PLP for machining center data

According to Cramer-von, the critical value of the test at 5% significance level is 0.22; therefore, the model can be used to estimate the failure trend of machining center.

5. Maintenance policy of machining center

5.1. Preventive maintenance

Machining center deteriorates with usage and can fail. If machining center fails, it would have a great effect on the product performance. In order to guarantee the reliability, appropriate maintenance should be paid on machining center. Actions to control (or reduce) equipment degradation are called PM and PM is classified into two groups – one is periodic PM and the other is sequential PM.

In order to improve the utilization of machining center, the users should develop the items of PM. There are generally two kinds of methods for PM of machining center: one is routine inspection, the other is regular inspection.

The goals of routine inspection of machining center are mainly used to examine whether there is enough lubricating oil, enough coolant liquid and whether the bolts, key connections and V-belt are loosened and whether there are leakage of oil, and so on. The routing testing items are shown in Table 11. There is regular inspection besides routine inspection for machining center. The regular inspection of machining center mainly includes spindle motor inspection, lubricate subsystem inspection, hydraulic subsystem, and so on. The regular inspection items are shown in Table 12.

5.2. Sequential preventive maintenance

As the parameter β=1.30, so the machining centers were in wearout life. That is to say the failure rate became higher with the increasing of the usage and maintenance times when the machining centers were in this life region. In order to improve the reliability, an appropriate maintenance policy should be optimized. Therefore, in this section we will select a sequential PM policy.

Considering the failure rate increasing over time, the failure rate between (i-1)th PM to i^{th} PM can be described in the Eq.(10)[16,17].

$$h_i(x) = \theta^{i-1}h(x+\varepsilon t_{i-1}), 1 \le \theta \le \mu, 0 \le \varepsilon \le 1 \quad (10)$$

Table 11. Routine inspection items

Num	Testing part	Testing items
1	Oil level gauge of the lubricate parts	• If there is enough oil • If the oil is contaminated
2	Surface of coolant liquid	• If amount of the coolant liquid is fit • If the coolant liquid is obvious contaminated • If the filter is clogged
3	Linear guide	• If there is enough lubricating oil • If the scratch chip board damages
4	Pressure gauge	• If the pressure is proper
5	V-belt	• If the tension is proper • If there are cracks and scratches
6	Pipe and appearance	• If there is the leakage of the oil • If there is the leakage of the coolant liquid
7	The moving parts	• If there are noise and vibrations • If the parts move smoothly
8	Panel	• If functions of the switch and handle are normal • If it displays alarm
9	Electric wire	• If there is disconnection • If the insulated coat is wearing out
10	Rotating part	• If there are noise and vibrations • If there is abnormal heat
11	Cleaning	• Clean the surface of the chuck, linear guide and chip machines
12	Workpiece	If the machining center keeps the machining accuracy under the control

Table 12. Regular inspection items

Num	Testing part		Testing items	Period
1	Hydraulic subsys-tem	Hydraulics Pipe joints	• Change the oil, clean the filters • Testing the leakage of the oil	6 mths 6 mths
2	Lubrication sub-system	Lubrication devices Pipe	• Clean the filters • Testing if there are the leakage, blockage and damage of pipes	1 year 6 mths
3	Cooling subsys-tems	Filter Chips plate	• Clean the chips plate • Change the coolant liquid, clean the filters and water tank	Depends on the situation
4	Pneumatic sub-system	Air filters	• Clean the air filters or change it	1 year
5	V-belt	Belt Pulley	• Test the tension • Clean the pulley	6 mths
6	Spindle motor	Sound, vibration and temperature rise	• Test the abnormal noise of the bearing • Clean the air filters	6 mths
7	Servo motor of X and Z axis	Sound and temperature rise	• Test the abnormal noise of the bearing and abnormal temperature rise	1 mth
8	Clamp subsystem	Clamp devices Cylinder	• Disassemble the clamp and clean it • Test the leakage of the cylinder	1 year 3 mths
9	Panel	Electrical devices Connection screws	• Test if there is odors, change color and damages of interface • Clean the connection screws	6 mths 1 mths
10	Electric subsystem	Limit switch Sensor Magnetic valve	• Test and fastening connection screws again • Test the function and activity of electric devices	6 mths 1 mths
11	X and Z axis	Clearance	• Measure the clearance by dial gage	6 mths
12	Base	Level of base	• Test and adjust the level of base by dial gage	1 year
13	Tool changer	Tool changer	• Test the origin of tool and adjust it	1mths

Where θ is increase factor of failure rate, ε is the repair factor of maintenance.

So the corresponding failure intensity function of PLP and the reliability function are

$$h_i(x)=\theta^{i-1}\lambda\beta(x+\varepsilon t_{i-1})^{\beta-1} \tag{11}$$

$$\begin{aligned} R_i(x)&=\exp[-\int_0^x h_i(x)dx] \\ &=\exp[-\lambda\beta\int_0^x \theta^{i-1}(x+\varepsilon t_{i-1})^{\beta-1}dx] \\ &=\begin{cases} \exp(-\lambda x^{\beta}), i=1; 0\le x\le T_1 \\ \exp\left\{-\dfrac{\lambda}{t_{i-1}}(\dfrac{\mu+1}{2})^{i-1}[\dfrac{(t_{i-1}+x)^{\beta+1}}{\beta+1}-\dfrac{t_{i-1}{}^{\beta+1}}{\beta+1}-\dfrac{x^{\beta+1}}{\beta+1}]\right\}, i=2,3,\ldots,N; 0\le x\le T_i \end{cases} \end{aligned} \tag{12}$$

When the reliability reduces to R_{min}, the sequential PM would be carried out. So from the above formula, we can get

$$\begin{cases} \exp(-xt^{\beta})=R_{\min}, i=1 \\ \exp\left\{-\dfrac{\lambda}{t_{i-1}}(\dfrac{\mu+1}{2})^{i-1}[\dfrac{(t_{i-1}+x)^{\beta+1}}{\beta+1}-\dfrac{t_{i-1}{}^{\beta+1}}{\beta+1}-\dfrac{x^{\beta+1}}{\beta+1}]\right\}=R_{\min}, i=2,3,\ldots,N \end{cases} \tag{13}$$

5.3. Maintenance cost

Denote the cost of repair by c_m, the cost of PM by c_p. the cost of replacement by c_r. Nakagawa [12] derived the following mean repair cost of N PM periods

$$C(T_i,N)=\frac{c_m\sum_{i=1}^{N}\int_0^{T_i}h_i(x)dx+(N-1)c_p+c_r}{\sum_{i=1}^{T_i}T_i} \tag{14}$$

The cumulative number of failures during the i^{th} interval of sequential PM is given by

$$F_i=\int_0^{T_i}h_i(x)dx=-\ln R_{\min} \tag{15}$$

Thus substituting Eq.(15) in Eq.(14) gets Eq.(16)

$$C(T_i,N)=\frac{(c_r-c_p)+N(c_p-c_m\ln R_{\min})}{\sum_{i=1}^{T_i}T_i}, i=1,2,\ldots,N \tag{16}$$

Our purpose is to seek both the optimal time T_i and number N which minimize $C(T_i,N)$ in Eq.(16). To find an N which minimizes $C(T_i,N)$, we form the inequalities

$$\begin{cases} C(T_i,N+1)\ge C(T_i,N) \\ C(T_i,N-1)\ge C(T_i,N) \end{cases} \tag{17}$$

Based on the empirical data, c_m=20000, c_p=10000, c_r=550000, R_{min}=0.7, substituting $\lambda=3.41\times10^{-5}$, β=1.3 in Eq.(16) and Eq.(17) gets the sequential PM periods and the mean cost C(N*) with 82.57. The results are shown in Table 13. The failure intensity function can be seen in Fig.6.

Table 13. Preventive maintenance period

i	T_i	t_i	*i*	T_i	t_i
1	1236	1236	13	384	8374
2	959	2195	14	362	8736
3	833	3029	15	341	9076
4	746	3775	16	321	9398
5	678	4453	17	303	9701
6	619	5072	18	286	9987
7	576	5648	19	271	10258
8	534	6182	20	256	10514
9	498	6680	21	242	10756
10	465	7145	22	229	10985
11	436	7581	23	217	11202
12	409	7990	24	206	11408

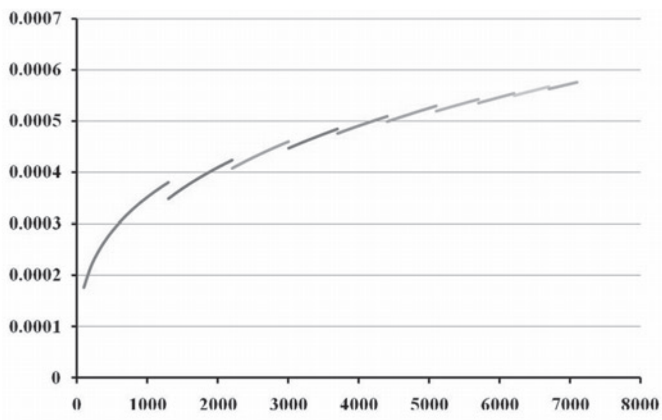


Fig.6. Failure intensity function k^{th} PM for machining center data

6. Conclusion

Synthetical design of reliability should be an integral part of design and management for the effective utilization of product. In this study the field failure data for 12 machining centers over five years were collected and analyzed. The following conclusions can be derived.

- The weakest subsystem of machining center is HS whose failures required the fourth longest repair time. It showed that the repairmen were not familiar with the HS, so the company should conduct repair training for the repairmen.
- The CA had the least failure modes and causes, so it was more likely to improve the reliability of CA. Therefore the manufacture factory of machining center should pay more attention to the design of CA.
- The failures of SS were mainly caused by the poor assembly and therefore the manufacture factory should do static balance test and dynamic equilibrium test to enhance the level of assembly.
- We have developed the PLP with $\lambda=3.41\times10^{-5}$ and β=1.30 for the machining center. It means the machining center is deteriorating with usage.
- Depending on the limitation of reliability and repair cost, the sequential PM policy was established with the mean cost C(24*) is equal to 82.57.

Finally, we should point out two implementation-related issues. The sequential PM policy is appropriate for the product that is deteriorating over time. The second issue deals with the different using conditions, where relevant PLP and maintenance policies need to be modified. It is an open issue for future study.

Acknowledgement

Our deepest gratitude goes first to the editor and reviewers for their constructive suggestions on the paper. And thank the authors of this paper's references whose work have contributed greatly to the completion of this thesis. Second, the authors would like to thank two Important National Science and Technology Specific Projects of China (2010ZX04014-011 and 2010ZX04014-016) and Graduate Innovation Research Program of JLU (No.20111057).

References

1. Ascher H, Feingold H. Repairable Systems Reliability. New York: Marcel Dekker Inc., 1984.
2. Cao JH, Cheng K. An Introduction to Reliability Mathematics. Beijing: Hinger Education Press, 2006.
3. Charles EE. An Introduction to Reliability and Maintainability Engineering. Beijing: TsingHua press and McGraw-Hill Education, 2008.
4. Deng LH, Kuang LJ. Performance Analysis of VDL-1000 Machining Center. Plant Maintenance Engineering 2006; 3, 34–35.
5. Hsu BM, Shu MH. Reliability assessment and replacement for machine tools under wear deterioration. International Journal of Advanced Manufacturing Technology 2010; 48, 355–365.
6. Jiang RY, Murthy DNP. Maintenance: Decision Models for Management. Beijing: Science Press, 2008.
7. Jiang RY, Guo Y. Two Simple Repairable Reliability Models. 2012 Reliability Technology Academic Conference of the National Machinery Industry, Changchun, 20-21 Aug 2012.
8. Larry HC. Evaluating the Reliability of Repairable Systems. 1990 Proceedings Annual Reliability and Maintainability Symposium, Los Angeles of USA, 23 -25 Jan 1990.
9. Larry HC. Reliability Growth Estimation With Missing Data-II. 1988 Proceedings Annual Reliability and Maintainability Symposium, Los Angeles of USA, 23 -25 Jan 1988.
10. Lindqvist BH. Statistical Modeling and Analysis of Repairable Systems. Proceedings of the 1st International Conference on Mathematical Methods in Reliability, Bucharest of Romania, 16-19 Sep, 1997.
11. Matthew MD. Excel 2010 the missing manual. Sebastopol: Pogue Press, 2010.
12. Nakagawa T. Periodic and sequential preventive maintenance policies. Journal of Applied Probability 1986; 23,536–542.
13. Wang YQ, Jia YZ, Yu JY, Yi SF. Field failure database of CNC lathes. International Journal of Quality and Reliability Management 1999; 16(4), 330–340.
14. Wang YQ, Jia YZ, Yu JY, Zheng YH, Yi SF. Failure probabilistic model of CNC lathes. Reliability Engineering and System Safety 1999; 65, 307–314.
15. Won JP, Yoon GK. Goodness-of-Fit tests for the Power-Law process. IEEE Transaction on Reliability 1992; 41(1), 107–111.
16. Wu SM, Derek CC. Preventive maintenance models with random maintenance quality. Reliability Engineering and System Safety 2005; 90, 99–105.
17. Zhou XJ, Xi LF, Lee J. Reliability-centered predictive maintenance scheduling for a continuously monitored system subject to degradation. Reliablity Engineering and System Safety 2007; 92, 530–534.

Prof. Zhao-jun YANG
Chuan-hai CHEN, Ph.D.
Associate Prof. Fei CHEN
College of Mechanical Science and Engineering
Jilin University
RenMin str., 5988 Changchun, China
e-mail: cchchina@foxmail.com, chench516@foxmail.com, chenfeicn@jlu.edu.cn

Qing-bo HAO
Department of Mechanics
Aviation University of Air Force
NanHu str., 2222 Changchun China
e-mail: haoqingbo@gmail.com

Bin-bin XU
College of Mechanical Science and Engineering
Jilin University
RenMin str., 5988 Changchun, China
e-mail: xubinbinjlu@foxmail.com

Article citation info:
GIRTLER J, ŚLĘZAK M. Four-state stochastic model of changes in the reliability states of a motor vehicle. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 156–160.

Jerzy GIRTLER
Marek ŚLĘZAK

FOUR-STATE STOCHASTIC MODEL OF CHANGES IN THE RELIABILITY STATES OF A MOTOR VEHICLE

MODEL STOCHASTYCZNY CZTEROSTANOWY ZMIAN STANÓW NIEZAWODNOŚCIOWYCH SAMOCHODU*

The properties of semi-Markov processes have been generally characterized and the applicability of the theory of such processes to the determining of the reliability of motor cars and other road vehicles has been explained. A formal description of the process of changes in the motor vehicle technical states considered as reliability states and a model of this process in the form of a one-dimensional stochastic process have been presented. The values of this process are the technical states of the motor vehicle in question that have significant practical importance. A four-state set of states interpreted as follows has been adopted: full (complete) serviceability, partial (incomplete) serviceability,task-limiting serviceability,and complete (total) unserviceability. Based on the initial distribution adopted and the functional matrix worked out, the boundary distribution of the process of changes in the technical (reliability) states of the motor vehicle has been defined. The probability of the vehicle being fully serviceable has been considered a measure of the vehicle reliability for a long period of vehicle operation. A possibility of defining the vehicle reliability in the form of a probability that a task would also be fulfilled by the vehicle being partially serviceable has also been indicated.

Keywords: *reliability, semi-Markov process, motor vehicle.*

W artykule scharakteryzowano ogólnie własności procesów semimarkowskich i uzasadniono możliwości ich zastosowania do określenia niezawodności samochodów i innych pojazdów drogowych. Przedstawiono formalny opis procesu zmian stanów technicznych samochodów uznanych za stany niezawodnościowe oraz model tego procesu w postaci jednowymiarowego procesu stochastycznego. Wartościami tego procesu są występujące w czasie eksploatacji stany techniczne samochodów, mające istotne znaczenie praktyczne. Przyjęto czterostanowy zbiór stanów o następującej interpretacji: stan zdatności pełnej (całkowitej), stan zdatności częściowej (niepełnej, niecałkowitej), stan niepełnej zdatności zadaniowej i stan niezdatności pełnej (całkowitej). Na podstawie przyjętego rozkładu początkowego i opracowanej macierzy funkcyjnej został określony rozkład graniczny procesu zmian stanów technicznych (niezawodnościowych) samochodu. Prawdopodobieństwo istnienia stanu zdatności pełnej (całkowitej) samochodu zostało uznane za miarę jego niezawodności w długim okresie czasu eksploatacji. Wskazano też na możliwość określenia niezawodności samochodu w formie prawdopodobieństwa, w którym uwzględniony został przypadek wykonania zadania przez samochód także wtedy, gdy znajduje się on w stanie zdatności częściowej.

Słowa kluczowe*: niezawodność, proces semi-Markowa, samochód.*

1. Introduction

In paper [12], a single-state reliability model of a passenger car was presented where one state of serviceability and ten states of unserviceability of the car were singled out, with the latter covering the cases where any of the major functional components of the car would fail. The following major functional components were selected there: 1) engine with fuel, lube oil, and coolant feeding systems; 2) clutch; 3) gearbox; 4) drive shaft; 5) driving axle; 6) steering and suspension system; 7) braking system; 8) electrical system; 9) body with chassis; and 10) measuring and monitoring equipment. An important good point of such a model is the fact that it reflects the serial reliability structure of the vehicle type considered and that it has arisen from the use of alternative classification of the vehicle reliability states into the state of serviceability s_0 and the states of unserviceability s_i (i = 1, 2, 3, …, 10), the latter constituting a set of unserviceability states S_n, i.e. $S_n = \{s_1, s_2, s_3, …, s_{10}\}$. However, in the practice of operation of motor vehicles, understood as both passenger cars and delivery vehicles, the states of partial serviceability, i.e. those intermediate between the full serviceability and unserviceability of the vehicle, may also be important; even the information whether the vehicle is fully serviceable may be of considerable value. In such a case, the approach to the reliability issue may be similar to that presented in paper [11], where the problem of building a model of the process of operation of diesel engines was addressed. The theory of semi-Markov processes was used to develop this model, too, as it was in the case described in [12]. This is important and, simultaneously, worth being emphasized here inasmuch as diesel engines are applied to some motor vehicles, both passenger cars and delivery truck or vans, and to other road transport facilities, e.g. buses etc. The process of occurrence of specific reliability states of motor vehicles is closely connected with the technical condition of the vehicles under consideration and it is one of the most important processes taking place during the vehicle operation stage. This process is composed of the technical states of a specific vehicle, following each other in succession and being causally connected together in time. It is obvious that the course of this process should be reasonable, i.e. it should be dictated by the optimizing criterion adopted, e.g. the expected value of the cost of operation of a vehicle of the specific type (which is important for passenger cars and delivery vehicles) or the any-time startability coefficient (which a top-priority parameter in the case of ambulances, fire-fighting vehicles, or police patrol cars). For the emergency vehicles mentioned here, a task may

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

be ordered at any time where immediate engine start and correct (reliable) functioning of all other equipment of such vehicles would be essential [8, 10]. Information about the reliability of motor vehicles of various types is critical for operating specific vehicles in a reasonable way. One of the most important indicators describing the reliability of motor vehicles is the probability that the vehicles will function correctly. For the solving of many problems related to the reliability of various devices, the theory of semi-Markov processes is successfully used more and more often. This theory may also be applicable when similar motor vehicle reliability problems are to be solved [2, 12]. Therefore, a semi-Markov reliability model of any motor vehicle with a four-state set of reliability states has been proposed here, with the interpretation of the reliability states having been presented in item 2 of this paper.

The semi-Markov processes are stochastic processes having special properties. Different definitions of such processes, with different ranges of generalization and exactitude, can be found in relevant publications. For the purposes of modelling the occurrence of specific technical states as reliability states of motor vehicles, the semi-Markov process (a family of random variables $\{W(t): t \geq 0\}$) may be defined with the use of the so-called homogenous Markov renewal process [1, 3, 13, 14, 17].

According to its definition, the semi-Markov process is a stochastic process with a discrete set of states and its realizations are constant functions in individual intervals (the functions having constant values within specific intervals of the operation time are random variables), continuous on the right. This definition also implies that this process is definite when its initial distribution $P_i = P\{Y(0) = s_i\}$ and functional matrix $\mathbf{Q(t)} = [Q_{ij}]$ are known, with the matrix being so defined that its elements are probabilities of transition from state s_i to state s_j within time not exceeding t ($i \neq j$; i, j =1, 2, …, k) and the probabilities are non-decreasing functions of variable t, which are denoted by $Q_{ij}(t)$ [4, 10].

A semi-Markov model of any real process may only be built when the states of this process can be so defined that the time of duration of the state existing at instant τ_n and the state obtainable at instant τ_{n+1} would not stochastically depend on the states that took place previously and on the time intervals of duration of those states.

The construction of a semi-Markov model $\{W(t): t \geq 0\}$ of a real process of changes in the technical states as reliability states of any motor vehicle is prerequisite for the applicability of the theory of semi-Markov processes. Such models are characterized by the following [5, 7, 8, 13, 14]:

1) The Markov condition is met, according to which the future evolution of the state of any motor vehicle as a research specimen (the process of changes in the reliability states during the vehicle operation stage) for which the semi-Markov model has been built should only depend on the vehicle state at the specific instant instead of the functioning of the same vehicle in the past, i.e. the *future* of the vehicle should exclusively depend on its *present* instead of its *past*.
2) The random variables T_i (representing the time of duration of state s_i regardless of the state to follow) and T_{ij} (representing the time of duration of state s_i providing that the next state of the process is s_j) have distributions other than exponential.

This means that the modelling aimed at designing a semi-Markov model of the process of changes in the technical states considered reliability states of motor vehicles should be done with taking into account an analysis of changes in the states of the real process, i.e. changes in the reliability states that occur during the stage of operation of the vehicles involved.

2. Formulation of the problem of reliability of a motor vehicle

The passenger car, like any other present-day road vehicle, is a complex technical system (Fig. 1), consisting of many elements having specific durability and reliability, which have been grouped into the functional components mentioned in the introduction [18].

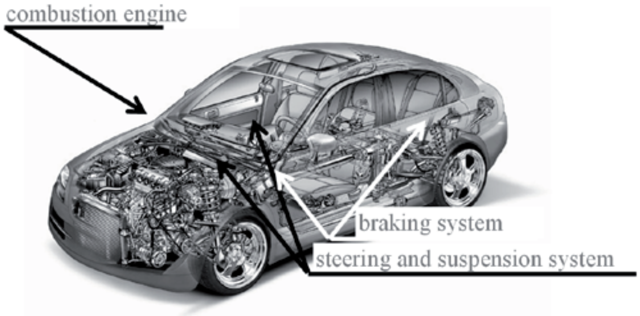


Fig. 1. Overall view of a passenger car, with indicated examples of the major functional components that are critical for the safety of car movement

The vehicle can function correctly (reliably) if all its elements, i.e. actually all its functional components, operate reliably when the vehicle is driven. However, such a situation during the vehicle operation stage is a probable rather than certain event. The probability of correct functioning of the vehicle decreases with time. The knowledge of this probability is important, especially after a longer time of the vehicle operation. When this probability is known, the preventive maintenance of the vehicle may be reasonably planned, because the knowledge of this probability and of the costs that might arise from a vehicle failure makes it possible to determine the risk of a failure to carry out a transport task. The probability of correct (reliable) functioning of a motor vehicle for a longer time of vehicle operation (theoretically at $t \rightarrow \infty$) may be relatively easily determined when a semi-Markov reliability model of the vehicle has been developed [5, 6, 9, 12, 13, 14]. Such a model represents changes in the technical states of the vehicle, which simultaneously are the reliability states of the vehicle.

3. The semi-Markov model of changes in the technical states considered reliability states of a motor vehicle

For every motor vehicle, as it is in the case of any diesel engine, the process of changes in its technical states is a process in which the time intervals of duration of every state are random variables. Individual realizations of these random variables depend on multiple factors, including the degree of wear of parts of the functional components of the said vehicles. For all the vehicle types, the wear of vehicle parts is weakly correlated with time [4, 7, 11, 16, 19]. This finding made it possible to forecast the technical state of the vehicles under consideration with ignoring the states that took place previously. This means that the semi-Markov process theory may be used to develop the reliability model of motor vehicles and thus, a more appropriate probabilistic mathematical model necessary for determining motor vehicle reliability indicators, especially the probability of correct functioning of the vehicles, may be obtained.

According to the deliberations presented in the introduction to this paper, the process of changes in motor vehicle reliability states $\{W(t): t \geq 0\}$ may be modelled by stochastic processes with a discrete set of states and with continuous time of duration of specific technical states of the vehicles. In mathematical terms, the models considered here as representing the process of changes in the technical state of motor vehicles, like in other technical objects, are functions that map the set of instants T into a set of technical states S. Therefore, for a model like this to be developed, a finite set of changes in the reliability (technical) states of the vehicles must be set up. Having adopted the usability of a motor vehicle for the carrying out of specific tasks as a criterion for the defining of separate states, we may differentiate the following set of classes (subsets) of the technical states, the classes being simply named "states" (that have significant importance in the vehicle operation practice), which simultaneously are vehicle reliability states [5, 7, 14]:

$$S = \{s_i;\ i = 1, 2, 3, 4\}. \tag{1}$$

Individual states s_i (i = 1, 2, 3, 4) being elements of set S should be interpreted as follows:

- s_1 – state of full (complete) serviceability, i.e. the technical state of any motor vehicle in which the vehicle may be used in the whole range of its capacity for which it was prepared at the designing and manufacturing stage;
- s_2 – state of partial (incomplete) serviceability, i.e. the technical state of a motor vehicle in which the vehicle may be used in the whole range of its capacity as it is in state s_1, but with significantly higher fuel consumption due to excessive engine wear or with increased braking distance due to wear of the brake mechanism (braking system) (Fig. 1);
- s_3 – state of task-limiting serviceability, i.e. the technical state of a motor vehicle in which the vehicle makes it possible to carry out only selected tasks, e.g. the state that has arisen from such a degree of wear of the combustion engine that higher vehicle speeds cannot be achieved (Fig. 1);
- s_4 – state of complete (total) unserviceability, i.e. the technical state of a motor vehicle in which the vehicle cannot be operated at all due to failure of the engine, brake mechanism (braking system), steering mechanism (steering system), suspension system, etc. (Fig. 1).

The state of complete or total unserviceability (s_4) is a result of a random event that is referred to as a total failure having occurred. Examples of such events may be [10, 18]: seizure of pistons in cylinders of a combustion engine, causing the pistons and crankshaft to be immobilized; burst of a brake hose, causing the brake fluid to leak out; deformation of a stub axle or a suspension arm, due to which the vehicle drive direction cannot be maintained; shearing failure of the key of an axle shaft spindle, due to which the road wheels cannot be driven, etc. The states of partial serviceability (s_2) or task-limiting serviceability (s_3) result from random events referred to as partial failures (or, often, minor defects or shortcomings), such as e.g. significant wear of injection equipment, pistons with piston rings, camshaft cams, thermostat failure, perforation of an exhaust silencer, fracture of a spring in the torsional vibration damper causing noisy operation of the clutch when disengaged, significant wear of the steering mechanism causing excessive steering play, etc.

Elements of set $S = \{s_i;\ i = 1, 2, 3, 4\}$ are values of process $\{W(t): t \geq 0\}$, which is composed of states $s_i \in S$ following one another in succession and being connected with each other by causality. A realization of such a process taken as an example has been presented in Fig. 2.

For motor vehicles, the differentiation between states $s_i \in S$ (i = 1, 2, 3, 4) is important inasmuch as it is essential for the vehicles to be only used when they are in state s_1 or, exceptionally, s_2. In the latter case, however, the vehicles should only be used for a time as short as possible and immediately after thatbe subjected to renovation.

This process is fully definite if its functional matrix is known [6, 9, 13]:

$$Q(t) = [Q_{ij}(t)], \tag{2}$$

with the non-zero elements of the matrix being interpreted as follows:

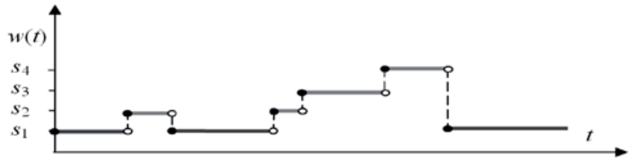


Fig. 2. An example of realizations of process $\{W(t): t \geq 0\}$ representing changes in the reliability states of a motor vehicle

$$Q_{ij}(t) = P\{W(\tau_{n+1}) = s_j,\ \tau_{n+1} - \tau_n < t \mid W(\tau_n) = s_i\};\ s_i, s_j \in S;\ i, j = 1, 2, 3, 4;\ i \neq j,$$

and when the initial distribution of this process is given:

$$p_i = P\{W(0) = s_i\},\ s_i \in S;\ i = 1, 2, 3, 4. \tag{3}$$

Depending on the strategy of maintaining the vehicles in a state that would enable them to carry out the tasks for which they were prepared at the designing and manufacturing stage, different variants of realizations of process $\{W(t): t \geq 0\}$ may be taken into account [11]. In the case of, especially, passenger cars and other vehicles prepared for the transportation of people, in consideration of the safety of motion of such vehicles, the most important variant is the one where the initial distribution of process $\{W(t): t \geq 0\}$ is as shown below:

$$p_1 = P\{W(0) = s_1\} = 1,\quad p_i = P\{W(0) = s_i\} = 0 \quad \text{for } i = 2, 3, 4 \tag{4}$$

and its functional matrix has the following form:

$$\mathbf{Q(t)} = \begin{bmatrix} 0 & Q_{12}(t) & 0 & 0 \\ Q_{21}(t) & 0 & Q_{23}(t) & 0 \\ Q_{31}(t) & 0 & 0 & Q_{34}(t) \\ Q_{41}(t) & 0 & 0 & 0 \end{bmatrix} \tag{5}$$

In this variant, an assumption is made that the vehicle operation is started when the vehicle is in the state of full serviceability (s_1). When the vehicle state changes into the state of partial serviceability (s_2), the vehicle is still used for the time sufficient to complete the transport task that has already been undertaken. The vehicle transition from state s_1 to state s_2 is a random event, which occurs with a probability of $p_{12}$when time T_{12}, which is a random variable, has elapsed. State s_2 lasts for time T_2, which is a random variable, too. When the current task is completed, the vehicle should be renovated by being subjected to an appropriate servicing procedure. Otherwise, when the vehicle use is continued, its technical state will change into the state of task-limiting serviceability (s_3), which may prevent the completion of the next task. The vehicle being in state s_3 should be renovated so that it is brought back to state s_1. The above means that a principle should be observed at the operation of motor vehicles to renovate the vehicles fully rather than partially. For this reason, the transition probabilities p_{32}, p_{42}, and p_{43} are zero (i.e. $p_{32} = 0$, $p_{42} = 0$, and $p_{43} = 0$), which has been taken into account in functional matrix (5) [11].

Functional matrix (5) represents changes in states $s_i \in S$ (i = 1, 2, 3, 4) of process $\{W(t): t \geq 0\}$. The matrix shows that these states can change as illustrated in the transition graph in Fig. 3.

According to the theory of semi-Markov processes [10, 13, 14], the probabilities of changes in the states of any technical object, i.e. of a motor vehicle as well, are defined by probabilities p_{ij} of transitions in the Markov chain $\{W(\tau_n): n = 0, 1, 2, ...\}$ inserted in process $\{W(t): t \geq 0\}$. These probabilities may be arranged in the following matrix of the probabilities of transitions:

$$P = [p_{ij};\ i,j = 1, 2, 3, 4] \tag{6}$$

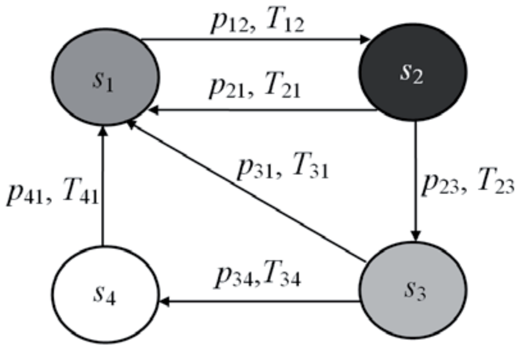


Fig. 3. A graph of changes in states $s_i \in S$ (i = 1, 2, 3, 4) of process $\{W(t): t \geq 0\}$

where: $p_{ij} = P\{W(\tau_{n+1}) = s_j \mid W(\tau_n) = s_i\} = \lim_{t\to\infty} Q_{ij}(t)$.

Matrix (6) makes it possible to determine the boundary distribution of process $\{W(t): t \geq 0\}$, the general interpretation of which is as given below:

$$P_j = \lim_{t\to\infty} P\{W(t) = s_j\} = \lim_{t\to\infty} P\{W(t) = s_j / W(0) = s_i\} \tag{7}$$

It follows from matrix (5) that this matrix has the form:

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ p_{21} & 0 & p_{23} & 0 \\ p_{31} & 0 & 0 & p_{34} \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{8}$$

From the theorem given in publication [14] on page 40, it is evident that a boundary distribution (7) of the process under consideration exists, which is defined by the formula [10, 12, 13]:

$$P_j = \frac{\pi_j E(T_j)}{\sum_{k=1}^{4} \pi_k E(T_k)}; \quad i = 1,\ 2,\ 3,\ 4 \tag{9}$$

with the boundary distribution π_j (j = 1, 2, 3, 4) of the Markov chain $\{W(\tau_n): n = 0, 1, 2, ...\}$,inserted in process $\{W(t): t \geq 0\}$, meeting the following equations:

$$\left.\begin{array}{l} [\pi 1, \pi 2, \pi 3, \pi 4] \cdot \begin{bmatrix} 0 & 1 & 0 & 0 \\ p_{21} & 0 & p_{23} & 0 \\ p_{31} & 0 & 0 & p_{34} \\ 1 & 0 & 0 & 0 \end{bmatrix} = [\pi 1, \pi 2, \pi 3, \pi 4] \\ \pi 1 + \pi 2 + \pi 3 + \pi 4 = 1 \end{array}\right\} \tag{10}$$

Having solved the system of equations (10) and with making use of relation (9), we will obtain the following formulas:

$$P_1 = E(T_1)M^{-1},\ P_2 = E(T_2)M^{-1},\ P_3 = p_{23}E(T_3)M^{-1},\ P_4 = p_{23}p_{34}E(T_4)M^{-1} \tag{11}$$

with: $M = E(T_1) + E(T_2) + p_{23}E(T_3) + p_{23}p_{34}E(T_4)$,

where:

$E(T_j)$ – expected value of the time of duration of state $s_j \in$ S (j = 1, 2, 3, 4);

p_{ij} – probability of the transition of process $\{W(t): t \geq 0\}$ from state s_i to state s_j (s_i, $s_j \in S$; i,j = 1, 2, 3, 4; $i \neq j$);

P_j – probability of process $\{W(t): t \geq 0\}$ being in state s_j (j = 1, 2, 3, 4).

Individual probabilities P_j (j = 1, 2, 3, 4) defined by formulas (11) should be interpreted as follows:

$$P_1 = \lim_{t\to\infty} P\{W(t) = s_1\},\ P_2 = \lim_{t\to\infty} P\{W(t) = s_2\},\ P_3 = \lim_{t\to\infty} P\{W(t) = s_3\},$$
$$P_4 = \lim_{t\to\infty} P\{W(t) = s_4\}$$

In the presented variant of changes in the differentiated reliability states of motor vehicles of any kind, the situations have also been taken into account where the user may risk thecarrying outof a task when the specific vehicle is in state s_2 (state of partial serviceability) or even to risk the completion of a task having been undertaken when the vehicle already is in state s_3 (task-limiting serviceability). The vehicle reliability may be measured by the probability of the vehicle being in state s_1, i.e. in the state in which it may be used in the whole range of its capacity, for a longer period of vehicle operation. In the situation where the transport task may be carried out by the vehicle being in state s_2, the vehicle reliability may be defined by probability $P = P_1 + P_2$. Probabilities P_3 and P_4 may, and should be, interpreted as the probabilities of an event that the vehicle would fail to carry out a task if it were previously in a state of long-time operation.

4. Recapitulation

The deliberations presented have shown that the process of changes in technical (reliability) states of motor vehicles is a stochastic process,which is discrete relative to the states and continuous relative to time,with four states interpreted as follows to be taken into account: s_1 – state of full (complete) serviceability, s_2 – state of partial (incomplete) serviceability, s_3 – state of task-limiting serviceability, and s_4 – state of complete (total) unserviceability. Actually, the technical state of every motor vehicle changes continuously; therefore, countable sets of vehicle states, i.e. sets consisting of an infinite number of elementary technical states, may be considered. The recognition of all the technical states of motor vehicles is neither possible nor advisable, for both technical and economic reasons. Hence, a need exists to divide the set of the technical states into a small number of classes (subsets) of the states. Having adopted the usability of a motor vehicle for the carrying out of specific tasks as a criterion to define separate states, we may differentiate, as mentioned before,certain classes (subsets) of the elementary technical states,with the classes being grouped into a set of states $S = \{s_1, s_2, s_3, s_4\}$, which may be considered a set of values of a stochastic process $\{W(t): t \in T\}$ whose realizations are constant in intervals and continuous on the right. Thus, this process is, in mathematical terms, a function mapping the set of instants T into a set of technical states S.

The semi-Markov processes used as models of real processes of operation of various technical objects are convenient tools in the research practice. They may also be suitable for the analysing of reliability of motor vehicles. This leads to a conclusion that the designing of a semi-Markov model of the process of operation of a technical object of any kind will provide a possibility of easy (thanks to the existing theory of semi-Markov processes) determining of probabilistic characteristics of motor vehicles.

The semi-Markov processes are more useful in practice as models of the real processes of changes in the states of technical objects than the Markov processes are. This is because the semi-Markov processes, where time is a continuous parameter and a finite set of states is considered, are characterized by the fact that the time intervals during which the processes remain in specific states are random variables with any distributions concentrated in the set $R_+ = [0, \infty)$. This feature distinguishes them from the Markov processes, where the intervals are random variables with exponential distributions.

An additional benefit of the use of semi-Markov processes (as it is when the Markov processes are used) is the fact that professional computer tools are available that make it possible to solve various systems of equations of states for such models of real processes.

References

1. Benjamin J.R, Cornell C.A. Rachunek prawdopodobieństwa, statystyka matematyczna i teoria decyzji dla inżynierów. Warszawa: WNT, 1977. Dane o oryginale: Probability, Statistics, and Decision for Civil Engineers. Copyright by McGraw-Hill, Inc., 1970
2. Cempel C, Natke H. G, Yao J.P.T. Symptom reliability and hazard for systems conditio motoring. Mechanical Systems and Signal Processing 2000; 14(3): 495–505.
3. Cinlar E. Markov reneval theory: a survey. Mag. Sciences 1975; 7: 727–752.
4. Gercbach I.B, Kordonski Ch.B. Modele niezawodnościowe obiektów technicznych. Warszawa: WNT, 1968.
5. Girtler J. Stochastyczny model procesu eksploatacji okrętowego silnika spalinowego. Zagadnienia Eksploatacji Maszyn 1989; 2: 79–88.
6. Girtler J. Możliwości zastosowania i przydatność procesów semimarkowskich jako modeli procesów eksploatacji maszyn. Zagadnienia Eksploatacji Maszyn 1996; 3: 419–428.
7. Girtler J. Diagnostyka jako warunek sterowania eksploatacją okrętowych silników spalinowych. Szczecin: WSM, 1997.
8. Girtler J.: Sterowanie procesem eksploatacji okrętowych silników spalinowych na podstawie diagnostycznego modelu decyzyjnego. Gdynia: AMW, 1989.
9. Girtler J. Physical aspects of application and usefulness of semi-Markovian processes for modeling the processes occurring in operational phase of technical objects. Polish Maritime Research 2004; 3: 7–12.
10. Girtler J. Operation of diesel engines as the index of their reliability and safety. Journal of KONES. International Combustion Engines 2003; 10(1-2): 17–24.
11. Girtler J. Semi-Markov model of diesel engines' operating process. The Archives of Automotive Engineering (Archiwum Motoryzacji) 2044; 3:165–196.
12. Girtler J, Ślęzak M. Application of the theory of semi-Markov processes to the development of a reliability model of an automotive vehicle. The Archives of Automotive Engineering (Archiwum Motoryzacji) 2012; 56(2): 15–27.
13. Grabski F. Teoria semi-markowskich procesów eksploatacji obiektów technicznych. Gdynia: AMW, 1982.
14. Grabski F. Semi-markowskie modele niezawodności i eksploatacji. Warszawa: IBS PAN, 2002.
15. Korczewski Z. Entropy function application In the selection process of diagnostic parameter of marine diesel and gas turbine engines. Polish Maritime Research 2010; 17(2): 29–35.
16. Niewczas A. Podstawy stochastycznego modelu zużywania poprzez tarcie w zagadnieniach trwałości elementów maszyn. Radom: Politechnika Radomska, 1989.
17. Pyke R. Markov reneval processes with finitely many states. Ann. Math. Statist. 1961; 32(4): 1243–1259.
18. Trzeciak K. Diagnostyka samochodów osobowych. Warszawa: WKiŁ, 1991.
19. Zwierzycki W. Prognozowanie niezawodności zużywających się elementów maszyn. Radom: ITE, 2000.

Prof. Jerzy GIRTLER, Ph.D. (Eng.)
Department of Ship Power Plants
The Gdansk University of Technology
ul. G. Narutowicza nr 11/12, 80-233 Gdańsk, Poland
e-mail: jgirtl@pg.gda.pl

Marek ŚLĘZAK, M.Sc. (Eng.)
Automotive Industry Institute
ul. Jagiellońska nr 55, 03-301 Warszawa, Poland
e-mail: m.slezak@pimot.org.pl

Article citation info:
AMANI N, ALI NM, MOHAMMED AH, SAMAT RA. Maintenance and management of wastewater system components using the condition index system, prediction processand costs estimation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 161–168.

Nima AMANI
Nasly Mohamed ALI
Abdul Hakim MOHAMMED
Roslida Abd SAMAT

MAINTENANCE AND MANAGEMENT OF WASTEWATER SYSTEM COMPONENTS USING THE CONDITION INDEX SYSTEM, PREDICTION PROCESS AND COSTS ESTIMATION

KONSERWACJA I ZARZĄDZANIE SYSTEMEM KANALIZACJI ŚCIEKOWEJ ZA POMOCĄ SYSTEMU WSKAŹNIKA STANU, PROCESU PRZEWIDYWANIA I SZACOWANIA KOSZTÓW

Component maintenance management of public building is complex and dynamic as the execution of the engineering management system is subjected to sensitive staff and users' requirements and high expectation of the top management for supporting the business. This paper presents the practices and survey need for maintaining the facilities systems in the building. The purpose of this study is maintenance time optimization of building component using the USACERL condition index (CI) system. To achieve this objective, cast iron pipe within wastewater plumbing system is surveyed using the financial analysis for implementation of optimal maintenance time based on limited cost. The findings show how a best time approach to plumbing system maintenance can assist the owner for decision making in component maintenance time based on existing cost.

Keywords: *maintenance management process, component, optimization, cost analysis, case study.*

Zarządzanie konserwacją instalacji budynku publicznego jest złożone i dynamiczne z uwagi na fakt, iż realizacja systemu zarządzania technicznego poddana jest zarówno wymaganiom personelu i użytkowników końcowych jak i oczekiwaniom kierownictwa w zakresie wsparcia rozwoju przedsiębiorstwa. W niniejszym artykule przedstawiono praktykę oraz badania dotyczące potrzeb wiążących się z konserwacją instalacji w budynkach. Celem tego opracowania jest optymalizacja czasu konserwacji tej części budynku za pomocą systemu wskaźnika stanu (condition index - CI) USACERL. Aby osiągnąć ten cel, przeanalizowano za pomocą analizy finansowej system kanalizacji ściekowej oparty na rurach żeliwnych pod kątem przyjęcia optymalnego czasu obsługi w oparciu o ograniczone koszty. Wyniki pokazują jak optymalne podejście czasowe do konserwacji systemu hydraulicznego może pomóc właścicielowi w procesie podejmowania decyzji w aspekcie czasu konserwacji na podstawie rzeczywistych kosztów.

Słowa kluczowe*: proces zarządzania konserwacją, instalacja, optymalizacja, analiza kosztów, studium przypadku.*

1. Introduction

Wastewater plumbing is an infrastructure system that has maintenance and establishment problems which are more serious than other installations such as electrical, communication, exterior system, etc. On the other hand plumbing system management is very important as it is related to health, customer's safety and environment protection [9]. A survey was done to identify the important building components and its degradation with respect to defects and indications, and its effect on clients, staffs and users of building. First, the top nine building facilities including interior surface, exterior surface, mechanical system, electrical system, communication system, clean water system, wastewater plumbing system, structural system and roofs were identified through the literature and through discussions with engineers and inspectors in the industry. Second, information that is related to the degradation of the building components was collected from a large owner organization, the Iranian Society of Consulting Engineers (ISCE). Comprehensive survey was then performed between experienced personnel at the ISCE in order to understand the various problems of components and statistics related to the level of difficulty in replacing, repairing, cleaning, and inspecting among components and facilities. Assessments show that mismanagement in maintenance of wastewater plumbing systems affects increasing of the building cost and decreasing of the wastewater network's service life [3]. From the other side, control and inspection of the whole pipeline is often very costly and time-consuming, and sometimes simply impossible [21]. Wastewater pipeline degradation results in uneconomical damages and high cost of replacement [9]. Therefore, condition assessment (CA) of plumbing in wastewater system is very essential and necessary for the prediction of proper repair time based on limited budget in relation to increasing the component service life and preventing early replacement, degradation and penalty costs [9, 21]. Due to the limited budget in maintenance management of wastewater plumbing system in Iran, hence it is uneconomical to replace the pipeline and piping [12]. This issue is very important on how to spend the limited financial allocation available for facilities and components maintenance to achieve the best return for their spending. There is a lack of integration between components maintenance time and cost decision making process in a building. The objective of this paper is to present a process to optimize the maintenance time and cost of cast iron pipe component in wastewater system of hotel buildings in Iran using the condition index (CI) systems. The paper presents a financial analysis using the data collected through financial and technical information. The technical information was collected from 30 installations consultants firms in Iran, while the financial information was collected from hotel's financial managers. Data is analyzed and simulated using the MS Excel software. This process by using the financial analysis assists in controlling the existing investment in facilities maintenance increasing components service life and, subsequently, preventing early deterioration and components replacement in public buildings.

2. Component condition index system

Systematic prediction by condition assessment method offers help to researchers in understanding the cost decision making in the best time for building facilities maintenance. Condition index (CI) presents the ability to form a basis for measuring rates of deterioration and prediction of condition for each component or facility [4]. The USACERL condition indexes were designed to support a purpose and quantitative means for component condition assessment while supplying a common language and explanation among users (assessor, engineer and inspector). The scale that is used in all of the USACERL indexes ranges from 0 to 100 and is divided into seven condition categories [10, 29]. The seven condition categories that set the arrangement of the index scale also need a guideline with the aim

Table. 1. USACERL condition index guide

Index	Category	Condition Description
86-100	Excellent	Very few defects. Component function is not impaired. No immediate work action is required, but routine or preventive maintenance could be scheduled for accomplishment.
71-85	Very Good	Minor deterioration. Component function is not impaired. No immediate work action is required but routine or preventive maintenance could be scheduled for accomplishment.
56-70	Good	Moderate deterioration. Component function may be somewhat impaired. Routine maintenance or minor repair may be required.
41-55	Fair	Significant deterioration. Component function is impaired but not seriously. Routine maintenance or minor repair is required.
26-40	Poor	Severe deterioration over a small percentage of the component. Less severe deterioration may be present in other portions of the component. Component function is seriously impaired. Major repair is required.
11-25	Very Poor	Critical deterioration has occurred over a large percentage or portion of the component. Less severe deterioration may be present in other portions of the component. Component is barely functional. Major repair or less than total reconstruction is required.
0-10	Failed	Extreme deterioration has occurred throughout nearly all or the entire component. Component is no longer functional. Major repair, complete restoration or total reconstruction is required.

to set the computed maintenance time for a component concerning the each index definition (condition description for each CI value). Table 1 presents these guidelines. It is very important that the guideline (condition description) displays the condition categories. This is because the use of definitions would influence integrated constraints on the formulation and the indexes for predicting maintenance time of components condition over time [10, 30].

Overtime, condition index (CI) moves from 100 to 0. When engineers install a component or material in a building, the condition index is 100 (excellent). Overtime, condition index for that component will reach below value 10. Basing on the definition of the CI scale, useful component failure happens when the CI falls around 10, which founds a functioning threshold limit for the model. For the unrepaired component lifecycle model, CI=10 when the time in service equals the expected service life. Hence, the profit of repair permits the deference of found rehabilitation required from component failure [30].

3. Statement of the Problem

Most important issue to successful component maintenance activities is a suitable cost allocated for a project. One of the reasons for change in facilities maintenance management and planning is due to limited allocation of costs [6]. Furthermore, lack of suitable costs allocation in a component maintenance work could affect the maintenance implementation [28]. Quality of maintenance works on the buildings components is dependent on the amount of costs allocation in this sector. Sufficient capital including staffing, inspecting, and financial is required for components maintenance works in the buildings [15]. Therefore, building managers or owners are responsible for management and allocation of maintenance costs for good maintenance outcome [18]. The statistics indicate increasing importance of maintenance and rehabilitation (M&R) of building components. Many components and facilities investment strategies lack of enough cost for components managers during its service life [19]. Hence, estimations and computations for maintenance costs planning and allocation is difficult and complex [1]. With respect to the statistics illustrated, it is an essential process in allocation of maintenance costs especially on the relative involvement of various components maintenance. Delay of some maintenance works accrued due to the cost allocated is limited and not sufficient to cover the requirement for building maintenance [7, 20]. Any decision making is based on existing costs and resources allocation of the buildings in maintenance activities [2]. Some researchers argue that decision making for component maintenance cost is synonymous with management [1]. Decision making for maintenance cost is a necessary process in management of building components and facilities [23].

4. Purpose of the Study

The main objective of this study is to develop a new process to optimize the maintenance time and cost of wastewater network of public building using the condition index system as a measurement method. The framework is focused on process to optimize building component maintenance time that has limited cost with respect to component condition assessment method (condition index tools) and economical management. This process involves the analysis of technical and financial information for a sample of five-star hotel buildings from the Iran.

5. The Case Study

Esteghlal Hotel with history of 50 years is one of the five-star hotels in Tehran. This hotel was found in 1962. Esteghlal Hotel is located in the down foot of the Alborz mountain range with 70,000 m^2 area, having two towers each with 15 floors, a total of 550 luxurious rooms and suites, not only it is the biggest and the most glorious hotel in the Tehran city, but also from various aspects it is most exclusive in Iran [8]. Table 2 shows the pipe dimension information for wastewater system.

Table. 2. Pipe dimension data for wastewater network of Esteghlal Hotel [24]

Diameter (inch)	2	3	4	5	6
Length (m)	1700	800	900	200	150

6. Research methodology

This section describes the research methodology adapted in the study. It explains all methods requested to achieve the objectives. To achieve the objectives of the research, the methodology is divided into three stages:

- Stage I – gathering the information,
- Stage II – developing a draft process,
- Stage III – verification of the process.

6.1. Stage I – gathering the information

The first step is gathering sufficient information on technical data related to the maintenance time of cast iron pipe from engineers' and inspectors' experiences over the past years with respect to moving the index from 100 to 10 during service life of the component. The rating sessions were carried out in small groups and at the normal work locations of the raters. The raters were first given general instructions by the researcher. This instruction is about the method of rating and determination of maintenance time for wastewater plumbing system during its service life. Each rater is then given a copy of the rating guidelines to use as rating cues (USACERL condition description), and a set of component rating sheets, given one at a time. As each rater completed a given sheet, it was collected by the researcher. After a given set of sheets was completed, the researcher reviewed the data during the session. Any rating and assessment of maintenance time that is different more than required standard deviations from the mean were flagged for a re-rate. This was done to allow raters the opportunity to correct certain ratings that may had been marked by mistake because of misunderstanding, distraction, misinterpretation or some other reason. The individual panel members ratings were averaged to obtain mean maintenance time for wastewater plumbing system. As depicted in Table 3, this predictive data is to achieve the information related to the maintenance time of cast iron pipe from stand point of engineers and inspectors' experiences in past years with respect to the moving the index from 100 to 0 during component's service life. The rating panel members that contributed to this development are consisted from contractors firms and related consultants on the wastewater plumbing system. The panel's opinion, indeed, represents a broad variety of experiences from commercial piping companies, installations firms and consulting firms. Their various position titles include directors and assistant directors of maintenance, piping foremen, piping inspectors, planners and estimators, civil engineers and installations engineers. As a group, the panel has experience regarding hot, cold, temperature, wet, and weather condition of related regions.

Table. 3. Gathering the information of technical data based on USACERL category (predictive data)

Index	Category	Maintenance year
100	Excellent	0
85	Very Good	9
70	Good	16.9
55	Fair	26.5
40	Poor	36.1
25	Very Poor	43.9
10	Failed	50

The financial information is related to the annual maintenance cost allocated for wastewater plumbing system in Esteghlal Hotel that is analyzed through gathering data by financial managers. The finan-

Table. 4. Financial information for wastewater system of Esteghlal Hotel

Historical data									
Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)
1990	≈0	1994	2500	1998	3900	2002	5500	2006	7000
1991	≈0	1995	2900	1999	4300	2003	6000	2007	7400
1992	≈0	1996	3100	2000	4800	2004	6200	2008	7900
1993	2000	1997	3300	2001	5000	2005	6700	2009	8200
Predictive data									
Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)	Year	Annual maintenance cost ($)
2011	≈0	2021	14600	2031	23800	2041	38800	2051	63300
2012	≈0	2022	15300	2032	25000	2042	40800	2052	66500
2013	≈0	2023	16100	2033	26300	2043	42800	2053	69800
2014	≈0	2024	16900	2034	27600	2044	45000	2054	73300
2015	≈0	2025	17800	2035	29000	2045	47200	2055	77000
2016	≈0	2026	18700	2036	30500	2046	49600	2056	80800
2017	≈0	2027	19600	2037	32000	2047	52100	2057	84900
2018	12600	2028	20600	2038	33600	2048	54700	2058	89100
2019	13200	2029	21600	2039	35200	2049	57400	2059	93600
2020	13900	2030	22700	2040	37000	2050	60300	2060	98200

Table. 5. Calculation of saving estimate for wastewater system of Esteghlal Hotel

Historical data			
Saving Condition Index	Maintenance Year/accidental	Computation basing on the Maintenance Year	Result ($)
Saving in Index 85	3.3	$\left(\sum_{i=1990}^{1992} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{1993})\times(\frac{4}{12})\right)$	666.7
Saving in Index 70	6.6	$\left(\sum_{i=1990}^{1995} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{1996})\times(\frac{6}{12})\right)$	8950
Saving in Index 55	9.9	$\left(\sum_{i=1990}^{1998} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{1999})\times(\frac{11}{12})\right)$	21641.7
Saving in Index 40	13.2	$\left(\sum_{i=1990}^{2002} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{2003})\times(\frac{2}{12})\right)$	38300
Saving in Index 25	16.5	$\left(\sum_{i=1990}^{2005} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{2006})\times(\frac{6}{12})\right)$	59700
Saving in Index 10	20	$\left(\sum_{i=1990}^{2009} \mathrm{FI}_i\right)$	86700
Predictive data			
Saving Condition Index	Maintenance Year	Computation basing on the Maintenance Year	Result ($)
Saving in Index 85	9	$\left(\sum_{i=2011}^{2019} \mathrm{FI}_i\right)$	25800
Saving in Index 70	16.9	$\left(\sum_{i=2011}^{2026} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{2027})\times(\frac{11}{12})\right)$	157000
Saving in Index 55	26.5	$\left(\sum_{i=2011}^{2036} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{2037})\times(\frac{6}{12})\right)$	401800
Saving in Index 40	36.1	$\left(\sum_{i=2011}^{2046} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{2047})\times(\frac{1}{12})\right)$	792100
Saving in Index 25	43.9	$\left(\sum_{i=2011}^{2053} \mathrm{FI}_i\right)+\left((\mathrm{FI}_{2054})\times(\frac{11}{12})\right)$	1279100
Saving in Index 10	50	$\left(\sum_{i=2011}^{2060} \mathrm{FI}_i\right)$	1808800

cial sheet was designed basing on the data collection method covering annual cost information for maintenance of wastewater piping system in Esteghlal Hotel. The financial information includes historical and predictive data. In this study, financial managers fill financial information from 1990 to 2009 for historical data (existing financial documents) and from 2011 to 2060 for predictive data. The historical data was collected for annual maintenance cost allocated at various condition index values for the cast iron pipe component in the wastewater plumbing system based on maintenance cost information in past 20 years (Table 4). The predictive data is selected basing on period of 50 years that corresponds to the useful lifespan of cast iron pipe which is approximately 50 years [16] using the prediction process and the average inflation rate computed from 1990 to 2009 (historical data) (Table 4). The average inflation rate of annual maintenance cost is 5% in Esteghlal Hotel. The information collected are stored in the saving sector for calculating the saving and investment ratio (SIR) for the condition index (from 100 to 10). The data, information and calculations are implemented basing on value of money and inflation rate computed for Iranian Rials currency (1 IRR = 0.0001 $).

6.2. Stage II – developing a draft process

The saving is total annual maintenance budget until maintenance time. The saving is calculated basing on the maintenance costs allocated for repair, service, inspection and clean annually in part of component maintenance until year *i*. The saving is estimated basing on USACERL condition index and maintenance year (technical information) for wastewater plumbing system in Esteghlal Hotel (Table 5) through the following formulas for historical and predictive data:

Saving formula for historical data:

$$\left(\sum_{i=1990}^{n} \mathrm{FI}_i\right) \tag{1}$$

Here, FI presents financial information of historical data, and n is Year-end of annual maintenance cost in desired index (index 85 to index 10).

Saving formula for predictive data:

$$\left(\sum_{i=2011}^{n} \mathrm{FI}_i\right) \tag{2}$$

Here, FI presents financial information of predictive data, and n is Year-end of annual maintenance cost in desired index (index 85 to index 10).

In this study the predictive data is important for implementation of maintenance prediction process. The predictive data is verified by historical data basing on the past 20 years. Second column of Table 5 (maintenance year – historical data) shows that the period of 20 years is divided into six parts basing on the USACERL condition index system. These data are accidental for verification of predictive data and prediction process during future years. The computation of saving estimate is done basing on the maintenance year in each condition index. For example in Table 5, in the first row of historical data (index 85) the saving is equal to sum of the financial information of historical data (Table 4) from 1990 to 1992 (3 years) plus 4/12 of 1993 (3 months) or in the third row of predictive data (index 55), the saving is equal to sum of the financial information of predictive data (Table 4) from 2011 to 2036 (26 years) plus 6/12 of 2037 (6 months).

Simulating economic analysis is carried out basing on the saving the investment ratio (SIR) for maintenance at various condition index values for the cast iron pipe component in the wastewater plumbing of Esteghlal Hotel. By using the simulation software the building managers are able to model and analyze the information without even knowing the complex mathematical models. Data required is collected through technical and financial information that is distributed among engineers, inspectors, and financial managers. Maintenance year is calculated from the technical data and saving is calculated from the financial data gathered. After data collection was completed, repair cost is computed. Repair cost is required cost for components restoration to excellent condition (operating period) after corrosion, broken, and other. Repair and maintenance cost is dependent on condition of weather, maintenance method, components functions, and management quality. Therefore, cost quantities have high standard deviation and there may not be any methods for accurate prediction of cost variation during future years [31]. Repair and maintenance cost of facilities and components is very low during first month (approximately 0). Function and lifespan enhancement result in increasing repair cost and repair cost is equal to replacement cost during final years. The market price fluctuation is a problem in relation to the accurate prediction of repair and maintenance cost. Therefore, accurate information and suitable statistics are very complex for computing component repair cost in field of component repair and maintenance [31]. The equations estimate repair costs as a percentage of the component purchase price (component replacement cost), so the equations should remain valid as long as the component purchase price goes up at the same rate as the cost of repairs [14]. The formulas for repair and maintenance costs estimate total accumulated repair costs based on accumulated hours of lifetime use. Repair and maintenance calculations are based on American Society of Engineers formulas [13]. There are other relevant studies in the field of repair cost computations including Sajadi and Moghadam in 2005 [26] and Means in 2008 [17]. In this study, the repair cost is computed basing on existing definition of repair cost and condition index method. This section defines the repair cost using the existing statistics of construction industry. The repair cost is analyzed by using the economic techniques and financial issues in repair and maintenance based on existing definitions. This equation is linear and uses the virtual variable:

$$C_r = C_p \times (Index\ a\ /\ Index\ b) \tag{3}$$

Table. 6. Pipe replacement cost C_p for wastewater system of Esteghlal Hotel

Historical data									
Year	Replacement cost ($)	Year	Replacement cost ($)	Year	Replacement cost ($)	Year	Replacement cost ($)	Year	Replacement cost ($)
1990	3805.3	1994	7890.8	1998	16362.5	2002	29088.9	2006	43954.5
1991	4566.4	1995	9469	1999	19635	2003	32321	2007	68177.5
1992	5479.7	1996	11362.8	2000	23562	2004	35349	2008	74995.3
1993	6575.7	1997	13635.4	2001	26180	2005	41073.5	2009	82494.8
Predictive data									
Year	Replacement cost ($)	Year	Replacement cost ($)	Year	Replacement cost ($)	Year	Replacement cost ($)	Year	Replacement cost ($)
2011	84800	2021	183076	2031	395249	2041	853313	2051	1842239
2012	91584	2022	197722	2032	426869	2042	921578	2052	1989618
2013	98910	2023	213540	2033	461018	2043	995304	2053	2148788
2014	106823	2024	230624	2034	497900	2044	1074929	2054	2320691
2015	115369	2025	249074	2035	537732	2045	1160923	2055	2506346
2016	124599	2026	268999	2036	580750	2046	1253797	2056	2706854
2017	134566	2027	290519	2037	627210	2047	1354100	2057	2923402
2018	145332	2028	313761	2038	677387	2048	1462429	2058	3157274
2019	156958	2029	338862	2039	731578	2049	1579423	2059	3409856
2020	169515	2030	365971	2040	790104	2050	1705777	2060	3682645

Table. 7. Repair cost C_r calculation for wastewater network of Esteghlal Hotel

Historical data				
Replacement cost ($)		Index/ Year	Calculation	Repair cost ($)
5479.7		85/ 3.3	5479.7× ((100-85) / (100-10))	876.7
9469		70/ 6.6	9469× ((100-70) / (100-10))	3124.7
16362.5		55/ 9.9	16362.5 × ((100-55) / (100-10))	8181.2
29088.9		40/ 13.2	29088.9 × ((100-40) / (100-10))	19198.6
41073.5		25/ 16.5	41073.5 × ((100-25) / (100-10))	34091
82494.8		10/ 20	82494.8 × ((100-10) / (100-10))	82494.8
Predictive data				
Replacement cost ($)	Index	Maintenance Year	Calculation	Repair cost ($)
156958	85	9	156958 × ((100-85) / (100-10))	25113
268999	70	16.9	268999 × ((100-70) / (100-10))	88769
580750	55	26.5	580750 × ((100-55) / (100-10))	290375
1253797	40	36.1	1253797 × ((100-40) / (100-10))	827506
2148788	25	43.9	2148788 × ((100-25) / (100-10))	1783494
3682645	10	50	3682645 × ((100-10) / (100-10))	3682645

Here, C_r presents repair cost in year *i*, and C_p is replacement cost in year *i*. *Index a* presents component condition in year *i*, and *Index b* is component condition in operating first year.

The unit replacement cost is according to the current price of cast iron pipe in the Iran's market. The replacement cost is calculated basing on dimension of wastewater plumbing system (size and length) of Esteghlal Hotel, price of cast iron pipe in the Iran's market, and average inflation rate for calculation of predictive data. The predictive data is calculated with inflation rate of 8% basing on average inflation rate of cast iron pipe in Iran's market from 2000 to 2010 (historical data) [11, 22]. The replacement cost is based on price index in Iran and including labor cost, transportation cost and the total cost of works [22]. The replacement cost was estimated through following formula and has been shown in Table 6:

$$C_p = L \times C \tag{4}$$

Here, C_p presents replacement cost in year *i, L* is lentgh of pipe, and *C* is cost basing on $/m.

Repair cost C_r is computed basing on standard equation 3. Between these condition index scales a parametric model of component repair cost is described as a comprehensive estimation of the corrective repair cost as a percentage of the total replacement cost in wastewater plumbing system of Esteghlal Hotel. Table 7 shows the repair costs in two situations of historical and predictive data.

In order to estimate the economic value of various component repair or replacement options, a saving to investment ratio (SIR) method is obtained. SIR is a numerical ratio and its size exhibits the economic execution of an investment. The SIR is saving divided by investment costs [25]. This model is selected due to: 1) each option results an individual ratio that shows the economic execution of that task action, 2) options can be used with various time horizons for comparing properly [5]. The SIR is illustrated by following equation:

$$SIR = Saving / Investment \tag{5}$$

Here, saving is total annual maintenance budget until repair time, and investment cost is repair cost in the year *i*. for a repair performance, the investment is the parametric evaluation of repair cost based on the condition index at year *i*. The saving is calculated based on the budget collected for maintenance annually in part of component maintenance until year *i*. Table 8 and 9 and Fig. 1 and 2 present the SIR for maintenance at various condition index values for the cast iron pipe wastewater system of Esteghlal Hotel in situations of historical and predictive data.

Table. 8. Financial analysis of optimum maintenance time based on USACERL condition index (historical data)

USACERL Index	Maintenance year	Investment estimation ($)	Saving estimation ($)	SIR
100	0	0	0	0
85	3.3	876.7	666.7	0.76
70	6.6	3124.7	8950	2.86
55	9.9	8181.2	21641.7	2.64
40	13.2	19198.6	38300	1.99
25	16.5	34091	59700	1.75
10	20	82494.8	86700	1.05

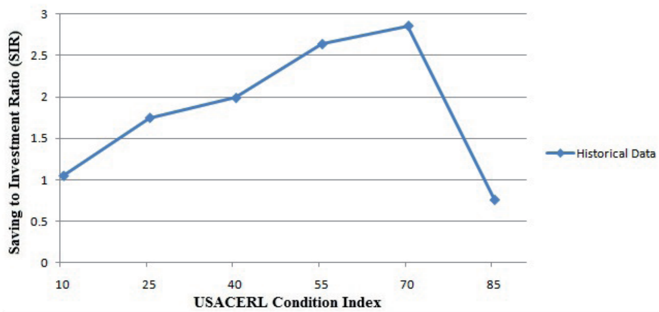


Fig. 1. Optimum maintenance time based on highest SIR (historical data)

Table 9 and Fig. 2 depict the analysis of optimum maintenance management of cast iron pipe in wastewater plumbing system for Esteghlal Hotel basing on a period of 50 years. The graph illustrates that when USACERL index reaches to 70, economic rate is high (SIR 1.76). Thus, the best time of maintenance occurs when USACERL CI is 70 with SIR 1.76. Building manager knows that the best decision for increasing cast iron pipe's service life based on existing maintenance cost is repair, cleaning, service, and renewal after 16 years (CI = 70 and SIR = 1.76). A ratio less than 1.0 indicates an uneconomic action

Table. 9. Financial analysis of optimum maintenance time based on USACERL condition index (predictive data)

USACERL Index	Maintenance year	Investment estimation ($)	Saving estimation ($)	SIR
100	0	0	0	0
85	9	25113	25800	1.02
70	16.9	88769	157000	1.76
55	26.5	290375	401800	1.38
40	36.1	827506	792100	0.95
25	43.9	1783494	1279100	0.71
10	50	3682645	1808800	0.49

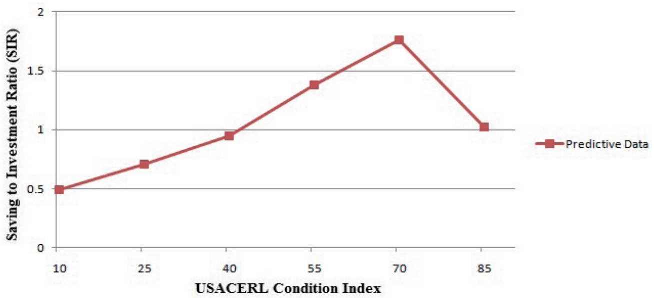


Fig. 2. Optimum maintenance time based on highest SIR (predictive data)

[25]. When the ratio is below 1.0, the economic efficiency of maintenance actions are nearly equal the replacement. Thus, if CI= 40, replacement close to the CI terminal value of 10 should be replaced. When CI reaches 40, maintenance time is 36 years basing on technical data. Thus, the wastewater system of Esteghlal Hotel should be replaced, no maintenance or repair after 36 years shall be done.

6.3. Stage III – verification of the process

Verification is the process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model. Verification is concerned on identifying and testing historical data in the model by comparing historical data and predictive data to analytical or highly accurate benchmark solutions [27]. Verification of process model is required when a predictive process is the end product. The predictive accuracy of the process must then reflect the strength of the inference being made from the historical database to the prediction. The verification of process model is motivated by the need for highly accurate process models for making predictions to support the maintenance management process model and by the current lack of guidelines, standards, and procedures for performing model. The verification assessment of

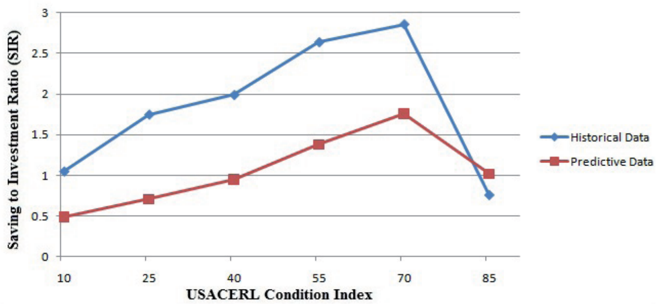


Fig. 3. Verification of process

model determines the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model. This information is used to decide whether or not the model has resulted in acceptable agreement with the experiment. The acceptable agreement decision focuses only on the level of match between the analysis outcome of historical data and predictive data. Verification is processes that collect evidence of a model's correctness or accuracy for a specific scenario; thus, verification cannot prove that a model is correct and accurate for all possible conditions and applications, but, rather, it can provide evidence that a model is sufficiently accurate.

Fig. 3 illustrates the financial analysis simulation of optimum index based on highest SIR in two situations of historical and predictive data. The results show that the highest SIRs occur when USACERL CI is 70 for two situations of historical and predictive data. The SIR systems soar from index 85 to index 70 and after that follow a downtrend from index 70 to index 10. The similarity of optimum index in two situations of historical and predictive data shows that this process is acceptable for prediction of maintenance time in wastewater system of Esteghlal Hotel.

7. Conclusion

This study presents the importance of survey of wastewater plumbing system using the USACERL condition index system in three stages including gathering the information, developing a draft process and verification of process for facilities correction maintenance in hotel buildings, in this case, cast iron pipe in the wastewater system of Esteghlal Hotel in Iran. This paper describes the development of a process for maintenance of building component using the USACERL condition index system. The systematic process was shown in a process framework for component maintenance for economical optimization for building component maintenance. This systematic process can provide a suitable decision making process on the maintenance time of building component based on existing budget for building owners and facilities managers.

Acknowledgement: *This study was supported by Universiti Teknologi Malaysia in the years 2010-2012 as a research project.*

References

1. Ali AS. Cost decision making in building maintenance practice in Malaysia. Journal of Facilities Management 2009; 7(4): 298–306.
2. Ali AS, Rahmat I, Hassan H. Involvement of key design participants in refurbishment design process. Facilities 2008; 26(9-10): 389–400.
3. Amani N, Nasly MA, Mohamed AH, Samat, R A. A survey on the implementation of facilities maintenance management system of buildings in Iran. Malaysian Journal of Civil Engineering 2012; 24(1): 85–95.
4. Amani N, Nasly MA, Samat RA. Infrastructure component assessment using the condition index system:literature review and discussion. Journal of Construction Engineering and Project Management 2012; 2(1): 27–34.
5. ASTM. Standard Practice for Measuring Benefit-to-Cost and Savings to Investment Ratios for Building and Building Systems, American Society for Testing and Materials, USA, 2002.
6. Boyle G. Design Project Management. NC Press, Burlington, 2003.
7. El-Haram MA, Horner MW. Factors affecting housing maintenance cost. Journal of Quality in Maintenance Engineering 2002; 8(2): 115–123.
8. Esteghlal Hotel. Persian Esteghlal International Co. http://esteghlalhotel.ir/, Iran, 2011.

9. Fahiminia M. Wastewater engineering in the earthquake-prone regions. Proceeding of 1the International Conference on Earthquake 2004; Kerman, Iran (in Persian).
10. Grussing MN. Building envelope life cycle condition evaluation using a distress-based methodology. Proceedings of the ASCE Specialty Conference Applications of Advanced Technology in Trans 2009: 2482–2490.
11. Index Mundi. Equivalent Data from the International Monetary Fund. http://www.indexmundi.com/iran/inflation_rate_(consumer_prices).html, Iran, 2011.
12. Khodabakhshi M. Plumbing system inspection based on API-570. Proceeding of the 1th International Conference on Technical Inspection and Non Distractive Test 2007; Tehran, Iran (in Persian).
13. Lazarus WF. Machinery Cost Estimates. University of Minnesota, 2012.
14. Lazarus WF. Estimating Farm Machinery Repair Costs. Extension Economist, University of Minnesota, 2008.
15. Lee HHY, Scott D. Overview of maintenance strategy, acceptable maintenance standard and resources from a building maintenance operation perspective. Journal of Building Appraisal 2009; 4(4): 269–78.
16. Lose Angeles HOA Management. Estimating Useful Life for Capital Assets. http://www.homeownersassociationmanagementla.com/Estimating-Useful-Lives-of-Building-Components.html, Lose Angeles, 1993.
17. Means RS. Square Foot Costs. 29th Annual Edition, RS Means Company Inc, 2008.
18. Mjema EAM. An analysis of personnel capacity requirement in the maintenance department by using simulation method. Journal of Quality in Maintenance 2002; 8(3): 253–73.
19. Mohd-Noor N, Hamid Y, Ghani AA, Haron SN. Building maintenance budget determination: an exploration study in the Malaysia government practice. Procedia Engineering 2011; 20: 435–444.
20. Narayan V. Effective Maintenance Management: Risk and Reliability Strategies for Optimizing Performance. Industrial Press, New York, 2003.
21. Pilch R, Szybka J, Broniec Z. Determining of hot water-pipe exploitation time on the basis of limiting states. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 203–207.
22. Price Index. President Deputy Strategic Planning and Control. http://www.spac.ir/Portal/Home/, Tehran, 2011.
23. Rahmat I. Managing Refurbishment Projects. UPENA, Universiti Teknologi MARA, Shah Alam, 2008.
24. Rashidi H. Interview, Department of Engineering, National Iranian Oil Company, Iran, 2011.
25. Ruegg TR, Marshall HE. Building Economic: Theory and Practice. Van Nostrand Reinhold- VNR Press, New York, 2003.
26. Sajadi SJ, Moghadam B. Cost assessment of maintenance and repair in industrial. Proceedings of the 3th Conference in Maintenance and Repair 2005; Tehran, Iran (in Persian).
27. Thacker BH, Doebling S W, Hemez F M, Anderson M C, Pepin J E, Rodriguez E A. Concepts of Model Verification and Validation. Los Alamos National Laboratory, California, USA, 2004.
28. Tilley PA, McFallen S L. Design and documentation quality survey designer's perspectives. CSIRO Press, Melbourne, 2000.
29. Uzarski DR. Development of Condition Indexes for Low Volume Railroad Track. U.S Army Construction Engineering Research Laboratories Press, Champaign, Illinois 1993.
30. Uzarski DR. Development of a track structure condition index (TSCI). Ph.D. Thesis, University of Illinois, Champaign, 1991.
31. Vafai MR, Mighani H, Borghai A M. Identification of mathematical model of maintenance and repair from Newholand machinery, TMISS Model. Journal of the Agriculture Modern Findings 2007; 2(2): 190–199.

Nima AMANI, Ph.D. Candidate
Faculty of Civil Engineering
Universiti Teknologi Malaysia
Skudai., 81310 UTM, Malaysia
e-mail: anima2@live.utm.my

Prof. Nasly Mohamed ALI
Faculty of Civil Engineering and Earth Resources
Universiti Malaysia Pahang
Pahang., 26300 Kuantan, Malaysia
e-mail: nasly@ump.edu.my

Prof. Abdul Hakim MOHAMMED
Faculty of Geoinformation and Real Estate
Universiti Teknologi Malaysia
Skudai., 81310 UTM, Malaysia
e-mail: abdhakim@utm.my

Dr. Roslida Abd SAMAT
Faculty of Civil Engineering
Universiti Teknologi Malaysia
Skudai., 81310 UTM, Malaysia
e-mail: roslida@utm.my

Article citation info:
MI J, LI Y, HUANG H-Z, LIU Y, ZHANG X-L. Reliability analysis of multi-state system with common cause failure based on bayesian networks. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 169–175.

Jinhua MI
Yanfeng LI
Hong-Zhong HUANG
Yu LIU
Xiao-Ling ZHANG

RELIABILITY ANALYSIS OF MULTI-STATE SYSTEM WITH COMMON CAUSE FAILURE BASED ON BAYESIAN NETWORKS

ANALIZA NIEZAWODNOŚCI SYSTEMU WIELOSTANOWEGO Z USZKODZENIEM SPOWODOWANYM WSPÓLNĄ PRZYCZYNĄ W OPARCIU O SIECI BAYESOWSKIE

Taking account of the influence of common cause failure (CCF) to system reliability and the widespread presence of multi-state system (MSS) in engineering practices, a method for reliability modeling and assessment of a multi-state system with common cause failure is proposed by taking the advantage of graphic representation and uncertainty reasoning of Bayesian Network (BN). The model is applied to a two-axis positioning mechanism transmission system to demonstrate its effectiveness and capability for directly calculating the system reliability on the basis of multi-state probabilities of components. Firstly, the reliability block diagram is built according to the hierarchy of structure and function of multi-state system. Then, the traditional Bayesian Networks model of the transmission system is constructed based on the reliability block diagram, failure logic between components and the failure probability distribution of them. In this paper, the β-factor model is used to analyze the CCF of the transmission system, and a new Bayesian network combining with CCF is established following by the implementation of reliability analysis. Finally, the comparison between the proposed method and the one without considering CCF is made to verify the efficiency and accuracy of the proposed method.

Keywords: *common cause failure (CCF), system reliability, multi-state system (MSS), Bayesian network (BN), β-factor model.*

Uwzględniając wpływ uszkodzeń spowodowanych wspólną przyczyną (CCF) na niezawodność systemów oraz powszechne występowanie w praktyce inżynierskiej systemów wielostanowych (MSS), zaproponowano metodę modelowania i oceny niezawodności systemu wielostanowego z uszkodzeniem spowodowanym wspólną przyczyną, która wykorzystuje reprezentację graficzną sieci Bayesa (BN) i oparte na nich wnioskowanie przybliżone. Model zastosowano do analizy układu przenoszenia napędu dwu-osiowego mechanizmu pozycjonowania. Zbadano w ten sposób skuteczność modelu oraz możliwość wykorzystania go do bezpośredniego obliczania niezawodności systemu na podstawie wielostanowych prawdopodobieństw elementów składowych. W pierwszej kolejności stworzono schemat blokowy niezawodności uwzględniający hierarchię struktury i funkcji badanego systemu wielostanowego. Następnie, w oparciu o schemat blokowy niezawodności, logikę uszkodzeń komponentów oraz rozkład prawdopodobieństwa uszkodzeń tych komponentów, skonstruowano tradycyjny model bayesowski układu przenoszenia napędu. W niniejszej pracy wykorzystano model współczynnika β do analizy CCF układu przenoszenia napędu oraz opracowano nową sieć Bayesa uwzględniającą CCF, po czym przeprowadzono na ich podstawie analizę niezawodności. Skuteczność i dokładność proponowanej metody sprawdzono poprzez porównanie jej z metodą nie wykorzystującą CCF.

Słowa kluczowe: *uszkodzenie spowodowane wspólną przyczyną (CCF), niezawodność systemu, system wielostanowy (MSS), sieci bayesowskie (BN), model współczynnika β.*

1. Introduction

Traditional reliability analysis is based on the assumption that events are binary, i.e., success and complete failure. However, different degrees of damage may have different effects on system performance. In this case, ignoring the partial failure of components and system may result in a huge error between the real system and the mathematical model used for system reliability analysis. Thus, Multi-State System (MSS) reliability analysis has received much attention over the past years [5, 10, 12, 14]. As a kind of complex system consisting of elements with different performance levels, multi-state system (MSS) widely exists in engineering practices, and Barlow and Wu et al. first introduced it in 1978 [5]. The basic concepts of MSS reliability were formulated and the system structure function was defined in Ref. [12]. Since the introduction of generalized multi-state k-out-of-n: G system in [9], many researches on multi-state k-out-of-n: G system modeling and optimization have been carried out [12, 13–16, 19, 21]. For instance, Liu and Kapur [19] established a reliability assessment model on dynamic multi-state non-repairable systems based on non-homogeneous Markov model, and determined a definition of two kinds of reliability measures for system performance. Ramirez-Marquez and Coit [21] presented and evaluated composite importance measures for MSS with multi-state components (MSMC), and a Monte Carlo simulation methodology was used to estimate the reliability of MSMC. An approach for the analysis of multi-state system with dependent elements was proposed by Levitin [14]. In order to analyze reliability of system having components with multiple failure modes, fault tree analysis method was incorporated into MSS by Huang [10]. The universal generating function (UGF) method for MSS reliability analysis was proposed by Ushakov [22], followed by an approach incorporating Markov model with UGF to calculate the reliability of MSS presented by Lisnianski and Levitin [12, 14, 15].

In the conventional system reliability analysis, it is often assumed that components and subsystems are independent. In reality, the common or shared fundamental cause, such as extreme environmental condition or design weaknesses may result in failure in multiple components, which is called Common Cause Failure (CCF). There are two fundamental methods to analyze the reliability of failure system with CCF: implicit method and explicit method [4]. Aiming at the CCF in power substation, the incorporated independent failure and the CCF was proposed in literature [6] by using the dynamic fault tree. Ref. [23] used an implicit method to incorporate CCF with a general procedure into system analysis to simplify the Boolean manipulation and quantification of fault trees. α-Factor model, multiple Greek letters model, and β-factor model etc. are used in the implicit method to estimate and evaluate the CCF in systems. A simplified α-factor model was implemented by Warren [24] to provide practical guidance for reducing complexity of the model, which needs to tailor to specific component failure criteria. BÖRCSÖK J gave details about the estimation and evaluation of common failures and used the β-factor model to assess a 1oo2 system in [4]. Levitin [15] utilized the universal generating function method to analyze system reliability by incorporating CCF into non-repairable MSS. An optimization model of MSS reliability in the presence of CCF was built and solved by Li et al. [13]. However, there is still lack of a general model that can be used to obtain reliability indexes of multi-state system with CCF [30].

Bayesian network (BN), which is based on a well-defined theory of probabilistic reasoning and the ability to express complex dependencies between random variables, has been applied to a variety of practical problems, especially in dependability assessment, risk analysis and maintenance [25]. Static BN was employed by Yin [28] and Zhou [30] to evaluate the reliability of MSS, after which many works have focused on the dynamic models by transforming the dynamic fault trees (DFT) into dynamic BN, the dynamic aspect had been integrated into system modeling and evaluating by Boudali and Dugan which can be referred to Ref. [1, 2]. To deal with the aforementioned issues in system reliability analysis, a multi-state system reliability analysis method is introduced in this paper by incorporating common cause failure into Bayesian network. The proposed method is applied to an engineering example of two-axis positioning mechanism of a satellite antenna.

The remainder of this paper is organized as follows: Section 2 briefly reviews the Bayesian network and introduces the procedure for system reliability CCF modeling. Bayesian network is applied to the multi-state transmission system in Section 3. Section 4 analyzes the reliability of MMS with CCF based on BN. It is demonstrated that BN is able to analyze the multi-state system reliability, and also deal with the common cause problems in multi-state systems.

2. Bayesian network and system reliability CCF modeling

2.1. Bayesian network

Bayesian network (BN), which was first proposed by Pearl [20], is a unique form of graphic description of probability relationships. A BN model typically includes a Directed Acyclic Graph (DAG) and a set of Conditional Probability Tables (CPT). In the directed acyclic graph, there are nodes representing random variables and directed arcs between pairs of nodes representing dependencies between them [7, 20].

Suppose A and B are two random events and the probability $P(B)>0$. Utilizing the Bayesian formula, the conditional probability of A given that B has happened is defined as follows:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

where $P(A)$ is the prior probability. Suppose that A has n possible states i.e. $a_1, a_2, \cdots, a_n$. Based on the total probability formula, $P(B)$ can be expressed as:

$$P(B) = \sum P(B|A = a_i)P(A = a_i) \tag{2}$$

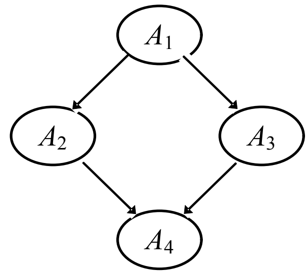


Fig. 1. A simple example of Bayesian network

A simple BN is shown in Figure 1. Here, A_i, i =1, 2, 3, 4 represents an event and the directed arc between any two nodes represents their causal relationships. A_1 does not have any incoming arcs and has no parents, thus it is a root node, which has marginal prior probability. A_2 and A_3 are intermediate nodes. Since A_4 is a node without outgoing arcs and having no children, it is called a leaf node. According to the chain rule of joint probability distribution, the joint probability of all the nodes in the aforementioned graph takes the following form [18],

$$P(A_1, A_2, A_3, A_4) = \prod_{i=1}^{4} P\{A_i | parent(A_i)\}$$
$$=P(A_1)P(A_2|A_1) \cdot P(A_3|A_1, A_2) \cdot P(A_4|A_1, A_2, A_3) \tag{3}$$

Due to the conditional dependence relationship of the events within the Bayesian network, it is convenient to derive the posterior probability from the prior probability as well as implementing reliability assessment of the system. The backward reasoning also makes it possible to evaluate the importance of components and carry out fault diagnosis. Because of the capability of describing the multi-state characteristics and the uncertainty of the logic relationship of events, Bayesian network is suitable for characterizing the random uncertainty and dependency of variables. Therefore, from its state description and inference mechanism, Bayesian network has been used in reliability analysis [7, 27, 28, 30, 31, 32].

2.2. Procedure for system reliability CCF modelling

The key to using BN to construct a CCF model for system reliability is to divide the failure rate λ_t of common cause component into an independent failure rate λ_I and CCF rate λ_c [8]. In this section, the procedure to construct a CCF model for system reliability based on BN will be shown by the modeling of several typical CCF systems. For components with m states, suppose state "1" means that the component is in good condition and state "0" means that the component fails. Between "0" and "1", there are m-2 states representing the different performance levels of the component.

(1) Series system

A series system with n components is the simplest and the most common model used in reliability analysis. Let $R_i(t)$ and $\lambda_i(t)$ denote the reliability and failure rate of component i, and $R_s(t)$ the reliability of the system. The mathematical model of the series system can be formulated as follows [26]:

$$R_s(t) = \prod_{i=1}^{n} R_i(t) = \prod_{i=1}^{n} e^{\int_0^t \lambda_i(t)dt} \tag{4}$$

The BN of a series system with two components (n = 2) considering CCF is shown in Figure 2. Here, X_i (i =1, 2) is a root node, D_1 and D_2 are intermediate nodes. X_i, C and D_1, D_2 are in two separate series.

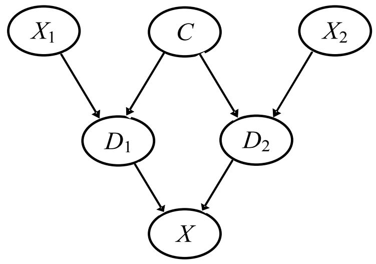


Fig. 2. The BN of series system considering CCF

In this paper, capital letters are used to express the events or the nodes of BN and small letters are variables representing the realization of the random events, so the mathematical expression of the reliability of the two-component series system is [27]

$$\begin{aligned} p(x=1) &= \sum_{x_1,x_2,c,d_1,d_2} p(x_1,c,x_2,d_1,d_2,x) \\ &= \sum_{d_1,d_2} \{p(x=1|d_1,d_2)\cdot \sum_{x_1,c}[p(d_1=1|x_1,c)p(x_1)p(c)] \\ &\cdot \sum_{x_2,c}[p(d_2=1|x_2,c)p(x_2)p(c)]\} \\ &= p(x_1=1)p(x_2=1)p(c=1) \end{aligned} \tag{5}$$

(2) Parallel system

The BN of a parallel system with two components is the same as that in a series system. The difference is that D_1 and D_2 shown in Figure 2 are now in parallel. The mathematical model of this two-component parallel system can be expressed as follows,

$$R_s(t) = 1 - \prod_{i=1}^{n}[1 - R_i(t)] \tag{6}$$

When $X_1 = X_2$, that is the two-components are identical, the expression of system reliability is,

$$\begin{aligned} p(x=1) &= \sum_{x_1,x_2,c,d_1,d_2} p(x_1,c,x_2,d_1,d_2,x) \\ &= \sum_{d_1,d_2} \{p(x=1|d_1,d_2)\cdot \sum_{x_1,c}[p(d_1=1|x_1,c)p(x_1)p(c)] \\ &\cdot \sum_{x_2,c}[p(d_2=1|x_2,c)p(x_2)p(c)]\} \\ &= 2p(x_1=1)p(c=1) - p^2(x_1=1)p(c=1) \end{aligned} \tag{7}$$

3. MSS reliability analysis based on BN

The requirements of high reliability and long lifetime for aerospace products are prevalent with the development of aerospace technology and have become the ultimate goal of the aerospace industry. As a commonly used satellite antenna control mechanism, the two-axis positioning mechanism has an important effect on the pointing accuracy of the antenna, to ensure the reliability of the launch and operation of the satellite [29]. This system has been widely used in military communications satellites, interplanetary exploration satellites, and earth observation satellites [11, 17, 29].

3.1. The two-axis positioning mechanism of the satellite antenna

As a key part for realizing a large range of satellite antenna rotation and high precision of positioning, the two-axis positioning mechanism is prone to failure, therefore, its reliability analysis is of great significance. According to its functions, the entire two-axis positioning mechanism can be divided into two subsystems: the transmission system and the control system. The transmission system achieves accurate positioning of the satellite antenna system through adjusting the direction of the pitch axis and the azimuth axis. Each axis is mainly composed of a motor, a reducer, and the shafts. According to the basic operating principle and the structure of the transmission system shown in Figure 3, the reliability block diagram can be expressed as a parallel-series structure. Each axis consists of stepper motor, drive shaft, and harmonic reducer that follow the series structure. In this paper, the pitch axis and the azimuth axis are simplified as parallel units, i.e. one axis failure does not cause the failure of the entire system.

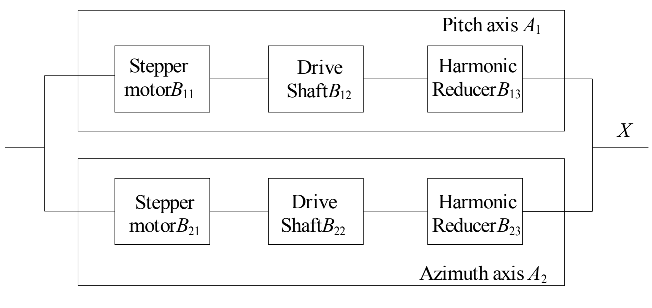


Fig. 3. The reliability block diagram of the transmission system

Since the pitch axis and the azimuth axis of the positioning mechanism are in a parallel structure, the relationship between the states of the two sets of components (A_1 is the state of the pitch axis, A_2 is the state of the azimuth axis) and the system state (X) is as follows: the failure of both sets of components will cause the failure of the whole transmission system. If one set of components is failed and the other is partially failed, the system is failed too. The system is working properly only when the two sets of components are working perfectly.

According to the descriptions above, we consider the two sets of components as two subsystems each with 3 states. Both the stepper motor and the harmonic reducer have 3 states: failure, partial failure, and working. The drive shaft has only 2 states: failure and working. We denote failure with X=0, partial failure with X=1, and working with X=2. The logic operator A_1, A_2 and X can be expressed as follows:

$$a_i = \min(b_{i1}, b_{i2}, b_{i3}) \quad (i=1,2) \tag{8}$$

$$x = \begin{cases} 0 & (a_1 a_2) \in ((00),(01),(10)) \\ 1 & others \\ 2 & (a_1 a_2) \in (22) \end{cases} \tag{9}$$

The Conversion of logic operators A_1, A_2 and X to conditional probability distributions can be obtained:

$$\begin{aligned} &P(a_i = 0 | b_{i1}, b_{i2}, b_{i3}) = 1 \quad b_{i1} \cdot b_{i2} \cdot b_{i3} = 0 \\ &P(a_i = 1 | b_{i1}, b_{i2}, b_{i3}) = 1 \quad others \quad i = 1,2 \\ &P(a_i = 2 | b_{i1}, b_{i2}, b_{i3}) = 1 \quad (b_{i1} b_{i2} b_{i3}) \in (222) \end{aligned} \tag{10}$$

$$\begin{aligned} &P(x = 0 | a_1, a_2) = 1 \quad (a_1 a_2) \in ((00),(01),(10)) \\ &P(x = 1 | a_1, a_2) = 1 \quad others \\ &P(x = 2 | a_1, a_2) = 1 \quad (a_1 a_2) \in (22) \end{aligned} \tag{11}$$

Since the azimuth axis has the same components as the pitch axis, each pitch axis is analyzed using the following multi-state reliability analysis approach.

3.2. Mapping reliability block diagram to Bayesian network

Because of the similarity between the principles of reliability block diagram and Bayesian network, the conversion algorithm of mapping reliability block diagram to Bayesian networks comprises the following steps: firstly, create a root node and assign its state space in the BN for each component of the system. Next, assign the prior probability to root nodes, and then create a corresponding node for each subsystem and assign its state space in the BN, and connect the nodes in the BN. Finally, assign the equivalent conditional probability distribution to the corresponding nodes according to its logic operator [30].

The mapping algorithm is performed on the example transmission system. The topology of the Bayesian network is shown in Figure 4. The conditional probability distribution of A_1 and A_2, which are converted from logic operator A_1 and A_2, are shown in the previous section 3.1.

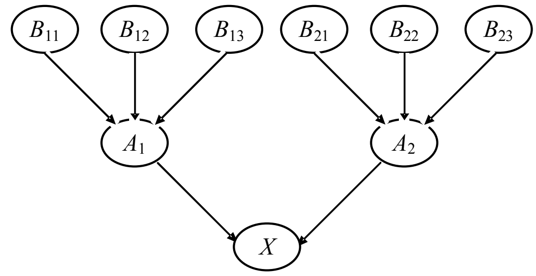


Fig. 4. Bayesian network of the transmission system

Due to the complexity of the two-axis positioning mechanism of satellite antenna and the shortage of data, the probability in each state of a root node are given based on experience [3, 4, 11]. For t = 3000h, the prior probability distribution of the root nodes of the Bayesian network is listed in Table 1 [11].

Table 1. Prior probability distribution of root nodes in Figure 4

Code	0	1	2
B_{i1}	3.0E_04	2.0E_04	9.995E_01
B_{i2}	7.0E_04	X	9.993 E_01
B_{i3}	6.0E_04	4.0 E_04	9.990E_01

3.3. Reliability analysis based on BN

The marginal probability distributions of an intermediate node and a leaf node that correspond to the probability distribution of the subsystems and the system can be obtained easily. In this section, according to equation (5) and (7), the probability of the system being in each state could be determined as follows:

$$p(X = j) = \sum p(b_{11}, b_{12}, b_{13}, a_1, b_{21}, b_{22}, b_{23}, a_2, x) \tag{12}$$

Where $b_{i1}, b_{i3}, a_1, a_2, x \in \{0,1,2\}$, $b_{i2} \in \{0,2\}$ and i=1,2; $j = 0,1,2$. Combining the former equation with the BN in Figure 4, the detailed expression of system reliability is,

$$\begin{aligned} p(X=2) &= \sum p(b_{11}, b_{12}, b_{13}, a_1, b_{21}, b_{22}, b_{23}, a_2, x) \\ &= \sum_{a_1, a_2} \{p(x=2 | a_1, a_2) \\ &\cdot \sum_{b_{11}, b_{12}, b_{13}} [p(a_1 | b_{11}, b_{12}, b_{13}) p(b_{11}) p(b_{12}) p(b_{13})] \\ &\cdot \sum_{b_{21}, b_{22}, b_{23}} [p(a_2 | b_{21}, b_{22}, b_{23}) p(b_{21}) p(b_{22}) p(b_{23})]\} \end{aligned} \tag{13}$$

Using the BNT and MATLAB to calculate the marginal probability distributions of A_1, A_2 and the probability of each state of the system X, the result is listed in Table 2 and Table 3.

Table 2. Marginal probability distributions of A_1 and A_2

The states of A_1 and A_2	0	1	2
Prob.	0.001599	0.000599	0.997802

Table 3. Probability of each state of transmission system

The states of *X*	0	1	2
Prob.	0.000004	0.004388	0.995608

Table 4. The probability of CCF occurs or not occurs

CCF	Occur (0)	Not occur (2)
Prob.	0.110932	0.889068

4. Reliability analysis of multi-state system with CCF based on BN

When taking hardware redundancy measures, the stability of system could be greatly improved while the other parts of the system will not be changed. Common cause failure has been regarded as a kind of important form of interrelated failure that is present in many mechanical components and systems. Being a major source of the failure of redundant systems, common cause failure increases the joint failure probability of each failure mode of the system and then leads to the reduction of the redundant system reliability.

Interrelated failures may occur to the mechanical components of an aerospace system and a nuclear system with high reliability requirement. The assumption that the failures of different parts of a system are independent, or the correlation of system failure is ignored, always leads to errors when analyzing system reliability [3, 4, 11, 17, 29].

4.1. β-factor model

Since it is difficult to measure the probability of common cause event accurately, the parametric modeling to assign the failure rate of components, such as α-factor model, β-factor model, has become the commonly used quantitative method. These parameter values are based on engineering experience and the published statistics of common cause failures. In this paper, the β-factor model is used to analyze the CCF of the transmission system [3].

Assume that Q_t is the total probability of failure for each component, which can be expanded into an independent contribution Q_I, and a dependent contribution Q_c. The parameter β is defined as the fraction of the total failure probability attributable to dependent failures [4]:

$$\begin{aligned}\beta &= \frac{Q_c}{Q_t} = \frac{Q_c}{Q_I + Q_c} = \frac{(1-\exp(-\lambda_c \cdot t))}{(1-\exp(-\lambda_t \cdot t))} \\ &= \frac{(1-\exp(-\lambda_c \cdot t))}{(1-\exp(-\lambda_I \cdot t))+(1-\exp(-\lambda_c \cdot t))}\end{aligned} \tag{14}$$

The range of the β-factor is from 0 to 0.25 (0 means no common cause failure) and reflects expert opinions about the β-factor in the range of 0.1% to 10% for hardware failure. If the associated components are sensitive to environmental stressors, it will have a high β-factor [4].

Assuming that the lifetime of all components obeys the exponential distribution, the failure rate of A_1 and A_2 can be computed following section 3 under $\lambda_1 = \lambda_2 = 0.002146$. The operating environment of the two-axis positioning mechanism of satellite antenna is complicated, so it has a high β-factor about 10%. According to equation (14), the dependent failure rate of common cause is $\lambda_c = 3.9194\text{E\_}05$. Table 4 shows the probability of common cause failure.

4.2. Reliability analysis of two-axis positioning mechanism transmission system with CCF based on BN

Using the proposed method in section 2 and section 3, we can generate the Bayesian network considering CCF, as shown in Figure 5.

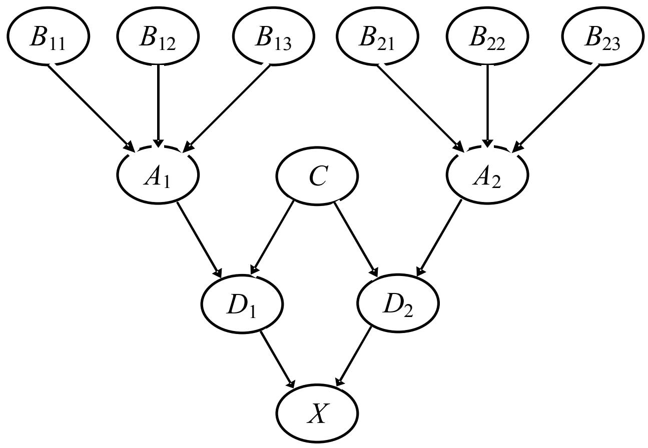


Fig. 5. The BN of transmission system considering CCF

The logic operator and the conditional probability distributions of D_1, D_2 and X are separately listed as the following equations,

$$d_i = \min(a_i, c) \quad (i=1,2) \tag{15}$$

$$x = \begin{cases} 0 & (d_1 d_2) \in ((00),(01),(10)) \\ 1 & others \\ 2 & (d_1 d_2) \in (22) \end{cases} \tag{16}$$

$$\begin{aligned} &P(d_i = 0 | a_i, c) = 1 \qquad others \\ &P(d_i = 1 | a_i, c) = 1 \qquad (a_i c) \in (12) \\ &P(d_i = 2 | a_i, c) = 1 \qquad (a_i c) \in (22) \end{aligned} \tag{17}$$

$$\begin{aligned} &P(x = 0 | d_1, d_2) = 1 \qquad (d_1 d_2) \in ((00),(01),(10)) \\ &P(x = 1 | d_1, d_2) = 1 \qquad others \\ &P(x = 2 | d_1, d_2) = 1 \qquad (d_1 d_2) \in (22) \end{aligned} \tag{18}$$

According to the equation (12) in section 3.3, the detailed expression of system reliability can be determined by equation (19). Finally, the probability of each state of system X considering CCF is listed in Table 5.

$$p(X=2)=\sum p(b_{11},b_{12},b_{13},a_1,d_1,b_{21},b_{22},b_{23},a_2,d_2,c,x)$$
$$=\sum_{d_1,d_2}\{p(x=2|d_1,d_2)$$
$$\cdot\sum_{a_1,c}\{p(d_1|a_1,c)\,p(c)\cdot\sum_{b_{11},b_{12},b_{13}}[p(a_1|b_{11},b_{12},b_{13})]p(b_{11})\,p(b_{12})\,p(b_{13})\}$$
$$\cdot\sum_{a_2,c}\{p(d_2|a_2,c)\,p(c)\cdot\sum_{b_{21},b_{22},b_{23}}[p(a_2|b_{21},b_{22},b_{23})]p(b_{21})\,p(b_{22})\,p(b_{23})\}\} \quad (19)$$

Table 5. Probability of each state of transmission system considering CCF

The states of *X*	0	1	2
Prob.	0.110934	0.003903	0.885163

From Table 5 we know that the probabilities for the transmission system to be in state 0, 1, 2 are $P(X=0)=0.110934$, $P(X=1)=0.003803$ and $P(X=2)=0.885163$. In addition, the reliability of the system is 0.885163 when $t = 3000$ h.

Comparing the data in Table 3 and Table 5, it is obviously that the probability for the transmission system in state 0 considering CCF in Table 5 is much greater than that in Table 3. That is, CCF has a remarkable effect on reliability of the transmission system in the two-axis positioning mechanism of the satellite antenna. So it is necessary to take measures to avoid or reduce the impact of common cause failure on the system.

5. Conclusion

In this paper, we proposed a Bayesian Network model to analyze the reliability of two-axis positioning mechanism transmission system used in the satellite antenna considering the CCF effects. The result shows that CCF has a considerable impact on the system reliability. This method can clearly express the influence of common cause failure on system reliability and does not require either the computation of minimal cut sets, or the determination of complicated algebraic expression of system unreliability.

In this paper, we only considered the cases where the component reliability states are three and the common cause failure only happened between the pitch axis and the azimuth axis. It should be pointed out that multi-state systems with more complex common cause failures and more states could still be handled using the proposed method. This is because the posterior probability distribution on each common cause and the importance of each component could be calculated conveniently under the Bayesian network framework. Furthermore, the proposed approach on qualitative analysis and quantitative assessment of multi-state system provided a reasonable guidance for system fault diagnosis and maintenance. How to avoid or reduce the impact caused by common cause failure will be investigated in our further work.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China under the contract number 51075061 and the National Programs for High Technology Research and Development of China under the contract number 2009018110019.

6. References

1. Boudali H, Dugan JB. A discrete-time Bayesian network reliability modeling and analysis framework. Reliability Engineering and System Safety 2005; 87: 337–349.
2. Boudali H, Dugan, JB. A continuous-time Bayesian network reliability modeling and analysis framework. IEEE Transaction on Reliability 2006; 55: 86–97.
3. Braband J, Hövel R v, Schäbe H. Probability of failure on demand-the why and the how. SAFECOMP 2009; LNCS 5775: 46–54.
4. Börcsök J, Schaefer S. Estimation and evaluation of common cause failures. Second International Conference on Systems 2007; 23: 867–882.
5. Barlow R, Wu AS. Coherent systems with multi-state elements. Journal Mathematics of Operations Research 1978; 3: 275–281.
6. Byeon Y, Kwon K, Kim J. Reliability analysis of power substation with common cause failure. International Conference on Power Engineering, Energy and Electrical Drives 2009; 467–472.
7. Ebrahimi A, Daemi T. Considering the rare events in construction of Bayesian network associated with power systems. PMAPS IEEE 2010; 659–663.
8. Ericson CA. Hazard analysis techniques for system safety. John Wiley & Sons, Inc, Canada, 2005.
9. Huang JS, Zuo MJ, Wu YH. Generalized multi-state k-out-of-n: G systems. IEEE Transactions on Reliability 2000; 49:105–111.
10. Huang X. Fault tree analysis method of a system having components of multiple failure modes. Microelectronics and Reliability 1983; 23: 325–328.
11. Jin N. Multi-state reliability analysis of a large deployable satellite antenna based on Bayesian networks. Xidian University, 2011.
12. Lisnianski A, Levitin G. Multi-state system reliability, assessment, optimization and application. World Scientific Publishing Co. Pte.Ttd, Singapore, 2003.
13. Li CY, Chen X, Yi XS. Reliability optimization of Multi-state system in presence of common cause failures. China Mechanical Engineering 2010; 21: 155–159.
14. Levitin G. A universal generating function approach for the analysis of multi-state systems with dependent elements. Reliability Engineering and System Safety 2004; 84: 285–292.
15. Levitin G. Incorporating common-cause failures into non-repairable multistate series-parallel system analysis. IEEE Transactions on Reliability 2001; 50: 380–388.
16. Li W, Zuo MJ. Reliability evaluation of multi-state weighted k-out-of-n systems. Reliability Engineering and Systems Safety 2008; 93: 160–167.
17. Liao Y, Li CJ, Feng XJ, Li M, Liao CW. Simulation and verification to the virtual prototype of two-axis-position structure based on Pro/E. Aerospace Shanghai 2006; 23: 59–65.
18. Liu Y, Singh C. Evaluation of hurricane impact on composite power system reliability considering common-cause failures. Int J Syst Assur Eng Manag 2010; 1: 135–145.
19. Liu Y W, Kapur KC. Reliability measures for dynamic multi-state nonrepairable systems and their applications to system performance evaluation. IIE Transactions 2006; 38: 511–520.
20. Pearl J. Fusion, propagation, and structuring in belief networks. Artificial Intelligence 1986; 29: 241–288.
21. Ramirez-Marquez JE, Coit D V. Composite Importance Measures for Multi-State Systems with Multi-State Components. IEEE Transactions and Reliability 2005; 54: 517–529.

22. Ushakov IA. Universal generating function. Sov. J. Comput. Syst. Sci.1986; 24: 118–129.
23. Vaurio JK. An implicit method for incorporating common-cause failure in system analysis. IEEE Transactions on Reliability 1998; 47: 173–180.
24. Warren CG. Common causes failures: Implementation of a simplified alpha factor model. Reliability and Maintainability Symposium 2010; 1–5.
25. Webern P, Medina-Oliva G, Simon C, Iung B. Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. Engineering Applications of Artificial Intelligence 2012; 25: 671–682.
26. Yin XW. Common cause failure model of system reliability based on Bayesian networks. International Journal of Performability Engineering 2010; 6: 255–268.
27. Yin XW, Qian WX, Xie LY. Common cause failure model of system reliability based on Bayesian networks. China Mechanical Engineering 2009; 20: 90–94.
28. Yin XW, Qian WX, Xie LY. Multi-state system reliability modeling and assessment based on Bayesian networks. Chinese Journal of Mechanical Engineering 2009; 45: 206–212.
29. Zhang H. System reliability analysis of dual-axis positioning mechanism of the satellite antenna. University of Electronic Science and Technology of China, 2011.
30. Zhou ZB, Jin G, Dong D, Zhou J. Reliability analysis of multistate systems based on Bayesian networks. 13th Annual IEEE International Conference and Workshop on Engineering of Computer Based Systems (ECBS 2006), 2006.
31. Zhou ZB, Li JJ, Li JQ, Xiao L. An implicit method for incorporating Bayesian networks into common cause analysis. World Congress on Global Optimization: Theory, Methods & Applications 2009; 12: 203–210.
32. Zhou ZB, Ma CQ, Zhou JL. Application of Bayesian networks reliability analysis of Binary-state system with multi-state components. Journal of Harbin Institute of Technology 2009; 41: 232–235.

Jinhua MI, Ph.D. candidate
Yanfeng LI, Ph.D. candidate
Prof. Hong-Zhong HUANG, Ph.D.
Associate Prof. Yu LIU, Ph.D.
Xiao-Ling ZHANG, Ph.D.
School of Mechanical, Electronic, and Industrial Engineering
University of Electronic Science and Technology of China
No. 2006, Xiyuan Avenue, West Hi-Tech Zone, Chengdu, Sichuan, 611731
P. R. China
e-mail: hzhuang@uestc.edu.cn

Article citation info:
KADYAN MS. Reliability and profit analysis of a single-unit system with preventive maintenance subject to maximum operation time. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 176–181.

Mukender Singh KADYAN

RELIABILITY AND PROFIT ANALYSIS OF A SINGLE-UNIT SYSTEM WITH PREVENTIVE MAINTENANCE SUBJECT TO MAXIMUM OPERATION TIME

ANALIZA NIEZAWODNOŚCI I ZYSKU DLA SYSTEMU JEDNOELEMENTOWEGO Z KONSERWACJĄ ZAPOBIEGAWCZĄ PODDANEGO MAKSYMALNEMU CZASOWI PRACY

This paper deals with the profit analysis of a reliability model for a single-unit system in which unit fails completely either directly from normal mode or via partial failure. The partially failed operating unit is shutdown after a maximum operation time for preventive maintenance. There is a single server who attends the system immediately whenever needed to conduct preventive maintenance at partial failure stage and repair at completely failure stage of the unit. The unit works as new after preventive maintenance and repair. The switch devices are considered as perfect. All random variables are assumed as independent and uncorrelated. The distribution of failure times, maximum operation time, preventive maintenance time and repair time are taken as general. Various reliability characteristics of interest are evaluated by using semi-Markov process and regenerative point technique. The tabular represantation of mean time to system failure (MTSF), availability and profit with respect to maximum rate of operation time has also been shown for a particular case.

Keywords: *single-unit system, reliability, preventive maintenance, maximum operation time, profit analysis.*

W niniejszej pracy przedstawiono analizę zysku modelu niezawodności dla systemu jednoelementowego, w którym element ulega całkowitemu uszkodzeniu bezpośrednio z trybu normalnego lub pośrednio na skutek częściowego uszkodzenia. Częściowo uszkodzona działająca jednostka jest wyłączana po upłynięciu maksymalnego czasu pracy w celu przeprowadzenia konserwacji zapobiegawczej. Pojedynczy serwer wspomaga bezzwłocznie system w momencie wystąpienia potrzeby przeprowadzenia konserwacji zapobiegawczej na etapie częściowego uszkodzenia oraz naprawy na etapie uszkodzenia całkowitego. Element działa jak nowy, po konserwacji zapobiegawczej i naprawie. Stan przełączników sieciowych uznaje się za doskonały. Wszystkie zmienne losowe traktowano jako niezależne i nieskorelowane. Rozkład czasów uszkodzeń, maksymalnego czasu pracy, czasu konserwacji zapobiegawczej i czasu naprawy przyjęto jako ogólne. Wybrane parametry niezawodnościowe oceniano za pomocą procesu semi-markowskiego i techniki odnowy RPT. Dla poszczególnych przykładów przedstawiono także tabelaryczne zestawienie średniego czasu do uszkodzenia systemu (MTSF), gotowości i zysku w odniesieniu do maksymalnego czasu pracy.

Słowa kluczowe*: system jednoelementowy, niezawodność, konserwacja zapobiegawcza, maksymalny czas pracy, analiza zysku.*

1. Introduction

Several researchers including Barlow and Larry [1], Nakagawa and Osaki [13], Murari and Goyal [12], Mokaddis et al. [11], Kumar et al. [6] and Renbin and Zaiming [14] have probed systems of one or more units making the assumption that the operating unit enters directly into the failed stage with constant failure rate and whenever the unit is under operation, it is continued until it fails.

But, in practice, there are many situations where a unit may fail completely either directly from normal mode or via various degraded stages. The devices subject to wear in reliability and the immune system of HIV infected individual in Bio-statistics can be considered as examples of such systems. However, continuous operation of a unit for a long time causes defects in the unit and increases the maintenance cost. Also, the continued operation and ageing of the systems gradually reduce their performance, reliability and safety. It can be seen from literature that preventive maintenance can slow the deterioration process of a repairable system and restore the system to a younger age or state. Therefore, preventive maintenance of the systems is necessary after a pre-specific period of time not only to maintain the operational power but may also reduce the failure and the degradation rate.

Keeping above facts in view, present paper deals with the cost-benefit analysis of a reliability model for a single-unit system in which unit fails completely either directly from normal mode or via partial failure. The partially failed operating unit is shutdown after a maximum operation time for preventive maintenance. There is a single server who attends the system immediately whenever needed to conduct preventive maintenance at partial failure stage and repair at completely failure stage of the unit. The unit works as new after preventive maintenance and repair. The switch devices are considered as perfect. All random variables are assumed as independent and uncorrelated. The distribution of failure times, maximum operation time, preventive maintenance time and repair time are taken as general. Various reliability characteristics for interest are evaluated by using semi-Markov process and regenerative point technique. The tabular represantation of MTSF, availability and profit with respect to maximum rate of operation time has also been shown for a particular case.

2. Notation

E	Set of regenerative states
O	The unit is operative and in normal mode
PFO	The unit is partially failed and operative
PFPm	The unit is partially failed and under preventive
FUr	The unit is failed and under repair
$f(t), F(t)$	Probability desity function (p.d.f.),Cumulative distribution function (c.d.f.) of the failure time from normal mode to complete failure
$f_1(t), F_1(t)$	p.d.f.,c.d.f. of the failure time from normal mode to partial failure
$f_2(t), F_2(t)$	p.d.f.,c.d.f. of the failure time from partial failure to complete failure
$g(t), G(t)$	p.d.f.,c.d.f. of the repair time of a failed unit
$z(t), Z(t)$	p.d.f.,c.d.f. of maximum operation time after partial failure
$h(t), H(t)$	p.d.f.,c.d.f.of the preventive maintenance time of the unit
*	Laplace transforms
©	Convolution

$E_0(t) = \overline{F(t)}\overline{F_1(t)}$ $\quad E_1(t) = \overline{Z(t)}\overline{F_2(t)}$

$E_2(t) = f_2(t)\overline{Z(t)}$ $\quad E_3(t) = z(t)\overline{F_2(t)}$

The system may be in one of the following states:

Up states $S_0(O)$, $S_1(PFO)$, $S_2(PFPm)$ Down states $S_3(FUr)$. Possible transitions between states along with cumulative distribution functions time are shown in Table 1.

Table 1.

From	S_0	S_1	S_2	S_3
S_0	-	$F_1(t)$	$F(t)$	-
S_1	-	-	$F_2(t)$	$Z(t)$
S_2	$G(t)$	-	-	-
S_3	$H(t)$	-	-	-

3. Reliability Analysis

Let $R_i(t)$ as the probability that the system survives during $(0, t) \mid E_0(t) = S_i$. To determine it we regard the failed states as absorbing state. The equations determining the reliability of the system. Hence we have:

$$R_0(t) = E_0(t) + f_1(t) © R_1(t)$$

$$R_1(t) = E_1(t) \tag{3.1}$$

By using Laplace transform technique, we can solve for $R_0^*(s)$ and is given by:

$$R_0^*(s) = E_0^*(s) + f_1^*(s)E_1^*(s) \tag{3.2}$$

The steady-state reliability of the system given by

$$R_0 = \lim_{s->0} sR_0^*(s) = \lim_{t->\infty} R_0(t) \tag{3.3}$$

4. Availability Analysis

Let $A_i(t)$ be the probability that the system is in upstate at instant *t* given that the system entered regenerative state *i* at *t=0*. The recursive relations for $A_i(t)$ are given by:

$$A_0(t) = E_0(t) + f_1(t) © A_1(t)$$

$$A_1(t) = E_1(t) + E_2(t) © A_2(t) + E_3(t) © A_3(t)$$

$$A_2(t) = g(t) © A_0(t)$$

$$A_3(t) = h(t) © A_0(t) \tag{4.1}$$

By taking Laplace transforms of the above equations and solving for $A_0^*(s)$, we get:

$$A_0^*(s) = \frac{N_1(s)}{D(s)} \tag{4.2}$$

where:

$$N_1(s) = E_0^*(s) + f_1^*(s)E_1^*(s)$$

$$D(s) = 1 - f_1^*(s)[h^*(s)E_3^*(s) + g^*(s)E_2^*(s)]$$

The steady-state availability of the system given by:

$$A_0 = \lim_{s->0} sA_0^*(s) = \lim_{t->\infty} A_0(t) \tag{4.3}$$

5. Busy Period of the Server due to Repair

Let $B_i^R(t)$ is defined as the probability that the system is busy due to repair at epoch t starting from state $S_i \in E$ .we have the following recursive relation:

$$B_0^R(t) = f_1(t) © B_1^R(t)$$

$$B_1^R(t) = E_2(t)©B_2^R(t) + E_3(t)©B_3^R(t)$$

$$B_2^R(t) = \overline{G(t)} + g(t)©B_0^R(t)$$

$$B_3^R(t) = h(t)©B_0^R(t) \quad (5.1)$$

By taking Laplace transforms of the above equations and solving for $B_0^{R*}(s)$, we get:

$$B_0^{R*}(s) = \frac{N_2(s)}{D(s)} \quad (5.2)$$

where: $N_2(s) = \overline{G^*(s)} f_1^*(s) E_2^*(s)$

$$D(s) = 1 - f_1^*(s)[h^*(s)E_3^*(s) + g^*(s)E_2^*(s)]$$

The steady-state of the busy period due to server is given by:

$$B_0^R = \lim_{s \to 0} sB_0^{R*}(s) = \lim_{t \to \infty} B_0^R(t) \quad (5.3)$$

6. Busy Period of the Server due to Preventive Maintenance

Let $B_i^P(t)$ is defined as the probability that the system is busy due to Preventive Maintenance at epoch t starting from state $S_i \in E$ .we have the following recursive relation:

$$B_0^P(t) = f_1(t)©B_1^P(t)$$

$$B_1^P(t) = E_2(t)©B_2^P(t) + E_3^P(t)©B_3^P(t)$$

$$B_2^P(t) = g(t)©B_0^P(t)$$

$$B_3^P(t) = \overline{H(t)} + h(t)©B_0^P(t) \quad (6.1)$$

By taking Laplace transforms of the above equations and solving for $B_0^{P*}(s)$, we get

$$B_0^{P*}(s) = \frac{N_3(s)}{D(s)} \quad (6.2)$$

where:

$$N_3(s) = \overline{H^*(s)}\, f_1^*(s)\, E_3^*(s)$$

$$D(s) = 1 - f_1^*(s)[h^*(s)E_3^*(s) + g^*(s)E_2^*(s)]$$

The steady-state of the busy period due to preventive maintenance server is given by:

$$B_0^P = \lim_{s \to 0} sB_0^{P*}(s) = \lim_{t \to \infty} B_0^P(t) \quad (6.3)$$

7. Expected Number of Visits by the Server

Let $N_i(t)$ be the expected number of visits by the server in (0,t] given that the system entered the regenerative state *i* at *t=0*. We have the following recursive relations for $N_i(t)$:

$$N_0(t) = f_1(t)©N_1(t)$$

$$N_1(t) = E_2(t)©[1 + N_2(t)] + E_3(t)©[1 + N_3(t)]$$

$$N_2(t) = g(t)©N_0(t)$$

$$N_3(t) = h(t)©N_0(t) \quad (7.1)$$

By taking Laplace transforms of the above equations and solving for $N_0^*(s)$, we get:

$$N_0^*(s) = \frac{N_4(s)}{D(s)} \quad (7.2)$$

where:

$$N_4(s) = f_1^*(s)[E_2^*(s) + E_3^*(s)]$$

$$D(s) = 1 - f_1^*(s)[h^*(s)E_3^*(s) + g^*(s)E_2^*(s)]$$

The steady-state of the busy period due o server is given by:

$$N_0 = \lim_{s \to 0} sN_0^*(s) = \lim_{t \to \infty} N_0(t) \quad (7.3)$$

8. Profit Analysis

Any manufacturing industry is basically a profit making organization and no organization can survive for long without minimum financial returns for its investment. There must be an optimal balance between the reliability aspect of a product and its cost. The major factors contributing to the total cost are availability, busy period of server and expected number of visits by the server. The cost of these individual items varies with reliability or mean time to system failure. In order to increase the reliability of the products, we would require a correspondingly high investment in the research and development activities. The production cost also would increase with the requirement of greater reliability.

The revenue and cost function lead to the profit function of a firm, as the profit is excess of revenue over the cost of production. The profit function in time t is given by:

P(t) = Expected revenue in (0, t] – Expected total cost in (0, t]

In general, the optimal policies can more easily be derived for an infinite time span or compared to a finite time span. The profit per unit time, in infinite time span is expressed as

$$\lim_{t \to \infty} \frac{P(t)}{t}$$

i.e. profit per unit time = total revenue per unit time – total cost per unit time. Considering the various costs, the profit equation is given as:

$$P = K_1 A_0 - K_2 B_0^R - K_3 B_0^P - K_4 N_0$$

where: K_1 = Revenue per unit up-time of the system,
K_2 = Cost per unit time for which server is busy in repair,
K_3 = Cost per unit time for which server is busy in preventive maintenance
K_4 = Cost per unit visit by the server.

9. Numerical Results

In this section,some of the results obtained for the above system are illustrated with a numerical example,we assume that

$$f(t) = \lambda e^{-\lambda t} \qquad f_1(t) = \lambda_1 e^{-\lambda_1 t} \qquad f_2(t) = \lambda_2 e^{-\lambda_2 t}$$

$$g(t) = \theta e^{-\theta t} \qquad h(t) = \beta e^{-\beta t} \qquad z(t) = \alpha e^{-\alpha t}$$

From equation (3.2), the time-dependent reliability is given by:

$$R_0^*(t) = \sum_{i=1}^{3} \frac{[s_i(s_i + \alpha + \lambda_2) + \lambda_1(2s_i + \alpha + \lambda + \lambda_1 + \lambda_2)]e^{s_i t}}{\prod_{j=1, i\neq j}^{3}(s_i - s_j)}$$

where $s_i (i = 1\ to\ 3)$ are the roots of the given equation.

$$s^3 + s^2(\alpha + \lambda + 2\lambda_1 + \lambda_2) + s(\lambda\alpha + \lambda\lambda_2 + 2\lambda_1\alpha + 2\lambda_1\lambda_2 + \lambda\lambda_1 + \lambda_1^2) + \\ + \lambda\lambda_1\alpha + \lambda\lambda_1\lambda_2 + \lambda_1^2\alpha + \lambda_1^2\lambda_2 = 0$$

Hence the mean time to failure of the system is calculated using the relation MTSF= $R_0^*(0) = \dfrac{(\alpha + \lambda_1 + \lambda_2 + \lambda_1)}{\lambda\alpha + \lambda\lambda_2 + \lambda_1\alpha + \lambda_1\lambda_2}$

Now from equation (4.2) the time-dependent availability of the system is given by:

$$A_0^*(t) = \sum_{i=1}^{5} \frac{[(s_i^2 + s_i(\alpha + \lambda_2 + 2\lambda_1) + \alpha\lambda_1 + \lambda_1^2 + \lambda_1\lambda_2 + \lambda\lambda_1)(s_i^2 + s_i\beta + \theta s_i + \theta\beta)]e^{s_i t}}{\prod_{j=1, j\neq i}^{5}(s_i - s_j)}$$

where $s_i (i = 1\ to\ 5)$ are the roots of the equation

$$\begin{aligned}
& s^5 + s^4(\lambda + 2\lambda_1 + \lambda_2 + \alpha + \beta + \theta) \\
& + s^3(\beta\alpha + \beta\lambda_2 + \alpha\theta + \lambda_2\theta + \beta\theta + 2\lambda_1\alpha + 2\lambda_1\lambda_2 + 2\lambda_1\beta + 2\lambda_1 + \lambda\alpha + \lambda\lambda_2 \\
& + \beta\lambda + \theta\lambda + \lambda_1\lambda + \lambda_1^2) + s^2(\beta\alpha\theta + \beta\theta\lambda_2 + 2\beta\theta\lambda_1 + 2\beta\lambda_1\lambda_2 + 2\alpha\theta\lambda_1 \\
& + \beta\alpha\lambda_1 + \theta\lambda_1\lambda_2 + \beta\alpha\lambda + \beta\lambda\lambda_2 + \alpha\lambda\theta + \lambda\theta\lambda_2 + \beta\theta\lambda + \lambda\alpha\lambda_1 + \lambda\lambda_1\lambda_2 + \beta\lambda\lambda_1 \\
& + \lambda_1\lambda\theta + \lambda_1^2\alpha + \lambda_1^2\lambda_2 + \beta\lambda_1^2 + \lambda_1^2\theta) + s(\beta\alpha\theta\lambda + \beta\lambda\theta\lambda_2 + \beta\lambda\lambda_1\lambda_2 \\
& + \alpha\theta\lambda\lambda_1 + \beta\lambda\theta\lambda_1 + \beta\alpha\theta\lambda_1 + \beta\theta\lambda_1\lambda_2 + \beta\lambda_1^2\lambda_2 + \alpha\theta\lambda_1^2 + \beta\theta\lambda_1^2) = 0
\end{aligned}$$

In case steady-state availability of the system given by

$$A_0 = \frac{\theta\beta(\alpha\lambda_1 + \lambda_1^2 + \lambda_1\lambda_2 + \lambda\lambda_1)}{\beta\theta(\alpha\lambda + \lambda\lambda_2 + \lambda\lambda_1 + \alpha\lambda_1 + \lambda_1\lambda_2 + \lambda_1^2) + (\beta\lambda_1^2\lambda_2 + \alpha\theta\lambda_1^2 + \beta\lambda\lambda_1\lambda_2 + \alpha\theta\lambda\lambda_1)}$$

From equation (5.2) the time-dependent busy period analysis due to server is given by:

$$B_0^{R*}(t) = \sum_{i=1}^{6} \frac{[\lambda_1\lambda_2(s_i^2 + s_i(\alpha + \lambda_2 + \beta) + \beta\alpha + \beta\lambda_2)]e^{s_i t}}{\prod_{j=1, i\neq j}^{6}(s_i - s_j)}$$

where $s_i (i = 1\ to\ 6)$ are the roots of the equation

$$\begin{aligned}
& s^6 + s^5(2\alpha + 2\lambda_2 + \beta + \theta + \lambda_1) \\
& + s^4(2\alpha\beta + 2\beta\lambda_2 + 2\alpha\theta + 2\theta\lambda_2 + \beta\theta + 2\alpha\lambda_1 + 2\lambda_1\lambda_2 + \beta\lambda_1 + \theta\lambda_1 + 3\alpha\lambda_2 \\
& + \lambda_2^2 + \alpha^2) + s^3(2\alpha\beta\theta + 2\beta\theta\lambda_2 + 2\beta\lambda_1\lambda_2 + 2\alpha\theta\lambda_1 + \beta\theta\lambda_1 + 3\alpha\beta\lambda_2 + \beta\lambda_2^2 \\
& + 3\alpha\theta\lambda_2 + 3\alpha\lambda_1\lambda_2 + \lambda_1\lambda_2^2 + \theta\lambda_1\lambda_2 + \beta\alpha^2 + \theta\alpha^2 + \alpha^2\lambda_1 + \alpha\beta\lambda_1 + \alpha^2\lambda_2 \\
& + \alpha\lambda_2^2 + \theta\lambda_2^2) + s^2(\beta\alpha\theta\lambda_2 + \beta\lambda_2^2\theta + \beta\lambda_1\lambda_2^2 + \alpha\theta\lambda_1\lambda_2 + \beta\theta\lambda_1\lambda_2 + \beta\alpha^2\theta \\
& + \beta\lambda_2\theta\alpha + \beta\lambda_1\lambda_2\alpha + \alpha^2\lambda_1\theta + \beta\lambda_1\theta\alpha + \beta\alpha^2\lambda_2 + \beta\lambda_2^2\alpha + \alpha^2\theta\lambda_2 + \lambda_2^2\alpha\theta \\
& + \alpha\beta\theta\lambda_2 + \lambda_1\lambda_2\alpha^2 + \lambda_1\lambda_2^2\alpha + \beta\lambda_1\lambda_2\alpha + \lambda_1\theta\lambda_2\alpha) \\
& + s(\beta\alpha^2\theta\lambda_2 + \beta\lambda_2^2\alpha\theta + \beta\lambda_1\lambda_2^2\alpha + \alpha^2\lambda_1\lambda_2\theta + \beta\lambda_1\lambda_2\alpha\theta) = 0
\end{aligned}$$

In case, Steady-state Busy period analysis due to server is given by

$$B_0^R = \frac{\beta\alpha\lambda_1 + \beta\lambda_1\lambda_2}{\alpha(\alpha\beta\theta + \beta\theta\lambda_2 + \beta\lambda_1\lambda_2 + \alpha\theta\lambda_1 + \beta\theta\lambda_1)}$$

From equation (6.2) the time-dependent busy period due to preventive maintenance of the system is given by:

$$B_0^{P*}(t) = \sum_{i=1}^{6} \frac{[\alpha\lambda_1(s_i^2 + s_i(\theta + \alpha + \lambda_2) + \alpha\theta + \theta\lambda_2)]e^{s_i t}}{\prod_{j=1, j\neq i}^{6}(s_i - s_j)}$$

where $s_i (i = 1\ to\ 6)$ are the roots of the equation

$$\begin{aligned}
& s^6 + s^5(2\alpha + 2\lambda_2 + \beta + \theta + \lambda_1) \\
& + s^4(2\alpha\beta + 2\beta\lambda_2 + 2\alpha\theta + 2\theta\lambda_2 + \beta\theta + 2\alpha\lambda_1 + 2\lambda_1\lambda_2 + \beta\lambda_1 + \theta\lambda_1 + 3\alpha\lambda_2 \\
& + \lambda_2^2 + \alpha^2) + s^3(2\alpha\beta\theta + 2\beta\theta\lambda_2 + 2\beta\lambda_1\lambda_2 + 2\alpha\theta\lambda_1 + \beta\theta\lambda_1 + 3\alpha\beta\lambda_2 + \beta\lambda_2^2 \\
& + 3\alpha\theta\lambda_2 + 3\alpha\lambda_1\lambda_2 + \lambda_1\lambda_2^2 + \theta\lambda_1\lambda_2 + \beta\alpha^2 + \theta\alpha^2 + \alpha^2\lambda_1 + \alpha\beta\lambda_1 + \alpha^2\lambda_2 \\
& + \alpha\lambda_2^2 + \theta\lambda_2^2) + s^2(\beta\alpha\theta\lambda_2 + \beta\lambda_2^2\theta + \beta\lambda_1\lambda_2^2 + \alpha\theta\lambda_1\lambda_2 + \beta\theta\lambda_1\lambda_2 + \beta\alpha^2\theta \\
& + \beta\lambda_2\theta\alpha + \beta\lambda_1\lambda_2\alpha + \alpha^2\lambda_1\theta + \beta\lambda_1\theta\alpha + \beta\alpha^2\lambda_2 + \beta\lambda_2^2\alpha + \alpha^2\theta\lambda_2 + \lambda_2^2\alpha\theta \\
& + \alpha\beta\theta\lambda_2 + \lambda_1\lambda_2\alpha^2 + \lambda_1\lambda_2^2\alpha + \beta\lambda_1\lambda_2\alpha + \lambda_1\theta\lambda_2\alpha) \\
& + s(\beta\alpha^2\theta\lambda_2 + \beta\lambda_2^2\alpha\theta + \beta\lambda_1\lambda_2^2\alpha + \alpha^2\lambda_1\lambda_2\theta + \beta\lambda_1\lambda_2\alpha\theta) = 0
\end{aligned}$$

Steady-state busy period analysis due to preventive maintenance is given by

$$B_0^P = \frac{\theta\alpha\lambda_1 + \theta\lambda_1\lambda_2}{\lambda_2(\alpha\beta\theta + \beta\theta\lambda_2 + \beta\lambda_1\lambda_2 + \alpha\theta\lambda_1 + \beta\theta\lambda_1)}$$

The time-dependent expected number of visits can be calculated from the equation (7.2) as

$$N_0^*(t) = \sum_{i=1}^{6} \frac{\{\lambda_1(\lambda_2 + \alpha)[(s_i + \theta)(s_i + \beta)(s_i + \alpha + \lambda_2)]\}e^{s_i t}}{\prod_{j=1, j\neq i}^{6}(s_i - s_j)}$$

where $s_i (i = 1\ to\ 6)$ are the roots of the equation

$$\begin{aligned}
& s^6 + s^5(2\alpha + 2\lambda_2 + \beta + \theta + \lambda_1) \\
& + s^4(2\alpha\beta + 2\beta\lambda_2 + 2\alpha\theta + 2\theta\lambda_2 + \beta\theta + 2\alpha\lambda_1 + 2\lambda_1\lambda_2 + \beta\lambda_1 + \theta\lambda_1 + 3\alpha\lambda_2 \\
& + \lambda_2^2 + \alpha^2) + s^3(2\alpha\beta\theta + 2\beta\theta\lambda_2 + 2\beta\lambda_1\lambda_2 + 2\alpha\theta\lambda_1 + \beta\theta\lambda_1 + 3\alpha\beta\lambda_2 + \beta\lambda_2^2 \\
& + 3\alpha\theta\lambda_2 + 3\alpha\lambda_1\lambda_2 + \lambda_1\lambda_2^2 + \theta\lambda_1\lambda_2 + \beta\alpha^2 + \theta\alpha^2 + \alpha^2\lambda_1 + \alpha\beta\lambda_1 + \alpha^2\lambda_2 \\
& + \alpha\lambda_2^2 + \theta\lambda_2^2) + s^2(\beta\alpha\theta\lambda_2 + \beta\lambda_2^2\theta + \beta\lambda_1\lambda_2^2 + \alpha\theta\lambda_1\lambda_2 + \beta\theta\lambda_1\lambda_2 + \beta\alpha^2\theta \\
& + \beta\lambda_2\theta\alpha + \beta\lambda_1\lambda_2\alpha + \alpha^2\lambda_1\theta + \beta\lambda_1\theta\alpha + \beta\alpha^2\lambda_2 + \beta\lambda_2^2\alpha + \alpha^2\theta\lambda_2 + \lambda_2^2\alpha\theta \\
& + \alpha\beta\theta\lambda_2 + \lambda_1\lambda_2\alpha^2 + \lambda_1\lambda_2^2\alpha + \beta\lambda_1\lambda_2\alpha + \lambda_1\theta\lambda_2\alpha) \\
& + s(\beta\alpha^2\theta\lambda_2 + \beta\lambda_2^2\alpha\theta + \beta\lambda_1\lambda_2^2\alpha + \alpha^2\lambda_1\lambda_2\theta + \beta\lambda_1\lambda_2\alpha\theta) = 0
\end{aligned}$$

The steady-state expected no of visit is given by:

$$N_0 = \frac{\theta\beta\lambda_1(\alpha+\lambda_2)^2}{\alpha\lambda_2(\alpha\beta\theta+\beta\theta\lambda_2+\beta\lambda_1\lambda_2+\alpha\theta\lambda_1+\beta\theta\lambda_1)}$$

10.Conclusion

The tabular behaviour of mean time to system failure (MTSF) with respect to maximum rate of operation time (α) is shown in table 2. It is observed that MTSF decrease with the increase of α. And, there is a further decline in their values when direct failure rate (λ) and partial failure rate (λ_1) increase. Tables 3 and 4 reflect respectively availability and profit of the system model decrease with the increase of maximum rate of operation (α), direct failure rate(λ) and partial failure rate (λ_1) for fixed values of other parameters. However, there is a substantial positive change in their values when repair rate (Θ) and preventive maintenance rate (β) increase. On the basis of the results obtained for a particular case it is analyzed that a system which undergoes preventive maintenance after a maximum operation time at partial failure stage can be made more profitable by increasing the repair rate of the system at its complete failure.

References

1. Barlow RB, Larry CH. Reliability analysis of a one-unit system. Operations Research Laboratories 1960; 9: 200–208.
2. Cox DR. Renewal theory, Chapman & Hall. 1962.
3. Gupta R. Probabilistic analysis of a two-unit cold standby system with two phase repair and preventive maintenance. Microelectronic Reliability 1986; 26: 13–18.
4. Khaled M. El-Said, Mohamed S. El-Sherbeny. Stochastic analysis of a two-unit cold standby system with two-stage repair and waiting time. Sankhya:The Indian Journal of Statistics 2010;72-B (1): 1–10.
5. Kumar J. Cost-Benefit Analysis of a Redundant System with Inspection and priority subject to degradation. International Journal of Computer Science Issues 2011; 8(6): 314–321.
6. Kumar J, Kadyan MS, Malik SC. Cost-Benefit Analysis of a two-unit parallel system subject to degradation after repair. Applied Mathematical Sciences 2010; 4(5): 2749–2758.
7. Kumar J, Kadyan MS, Malik SC. Cost Analysis of a Two-Unit Cold Standby System Subject to Degradation, Inspection and Priority. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14(4): 278–283.
8. Kowada M, Bando A. Preventive maintenance of a one-unit system with two types of repair. Microelectronic Reliability 1982; 22: 287–293.
9. Mahmoud MAW, Moshref ME. On a two-unit cold standby system considering hardware, human error failures and preventive maintenance. Mathematical and Computer Modelling 2010; 51: 736–745.
10. Malik SC, Nandal P, Barak MS. Reliability analysis of a system under preventive maintenance. Journal of Mathematics and System Sciences 2009; 5(1):92–115.
11. Medhi J. Stochastic Processes, Wiley Eastern Limited, India. 1982.
12. Mokaddis GS, Labib SW, Ahmed AM. Analysis of a two-unit warm standby system subject to degradation, Microelectron. Reliab. 1997; 37(4): 641–647.
13. Murari K, Goyal V. Comparison of two unit cold standby reliability models with three types of repair facilities. Microelectron. Reliab. 1984; 24(1): 35–49.
14. Nakagawa T, Osaki S. Reliability analysis of a one-unit system with unrepairable spare units and its optimization applications. Quarterly Operations Research (1976); 27(1): 101–110.
15. Renbin Liu, Zaiming Liu. Reliability analysis of a one-unit system with finite vacations, Management Science Industrial Engineering (MSIE). International Conference 2011; 248–252.
16. Singh SK, Agrafiotis GK. Stochastic analysis of a two-unit cold standby system subject to maximum operation and repair time. Microelectronic Reliability. 1995; 35(12): 1489–1493.

Table 2.

α ↓	Mean Time to System Failure(MTSF)		
	$\lambda=.13,\lambda_1=.17, \lambda_2=.21,\theta=2.1, \beta=2.7$	$\lambda=.16,\lambda_1=.17, \lambda_2=.21,\theta=2.1, \beta=2.7$	$\lambda=.13,\lambda_1=.20, \lambda_2=.21,\theta=2.1, \beta=2.7$
5	3.550864	3.228058	3.262956
10	3.444336	3.131214	3.149022
15	3.407846	3.098042	3.109995
20	3.389411	3.081283	3.090279
25	3.378289	3.071172	3.078384
30	3.370849	3.064408	3.070426
35	3.365521	3.059565	3.064729
40	3.361519	3.055926	3.060448
45	3.358402	3.053092	3.057114
50	3.355905	3.050823	3.054444

Table 3.

α ↓	Availability				
	λ=.13,λ_1=.17, λ_2=.21,θ=2.1, β=2.7	λ=.16,λ_1=.17, λ_2=.21,θ=2.1, β=2.7	λ=.13,λ_1=.20, λ_2=.21,θ=2.1, β=2.7	λ=.13,λ_1=.17, λ_2=.21,θ=2.6, β=2.7	λ=.13,λ_1=.17, λ_2=.21,θ=2.1, β=3.7
5	0.891564	0.880701	0.883569	0.901317	0.904315
10	0.890324	0.879323	0.881995	0.899959	0.903512
15	0.889891	0.878842	0.881444	0.899485	0.903231
20	0.889671	0.878597	0.881163	0.899244	0.903088
25	0.889537	0.878449	0.880992	0.899098	0.903001
30	0.889448	0.878349	0.880878	0.899	0.902943
35	0.889383	0.878278	0.880796	0.898929	0.902902
40	0.889335	0.878224	0.880735	0.898876	0.90287
45	0.889297	0.878183	0.880687	0.898835	0.902846
50	0.889267	0.878149	0.880648	0.898802	0.902826

Table. 4

α ↓	Profit				
	λ=.13,λ_1=.17, λ_2=.21,θ=2.1, β=2.7,K_1=5000, K_2=150,K_3=75, K_4=50	λ=.16,λ_1=.17, λ_2=.21,θ=2.1, β=2.7,K_1=5000, K_2=150,K_3=75, K_4=50	λ=.13,λ_1=.20, λ_2=.21,θ=2.1, β=2.7,K_1=5000, K_2=150,K_3=75, K_4=50	λ=.13,λ_1=.17, λ_2=.21,θ=2.6, β=2.7,K_1=5000, K_2=150,K_3=75, K_4=50	λ=.13,λ_1=.17, λ_2=.21,θ=2.1, β=3.7,K_1=5000, K_2=150,K_3=75, K_4=50
5	4432.517	4375.405	4390.878	4482.644	4496.983
10	4426.083	4368.243	4382.708	4475.604	4492.754
15	4423.836	4365.743	4379.847	4473.147	4491.276
20	4422.693	4364.471	4378.389	4471.896	4490.524
25	4422	4363.7	4377.505	4471.138	4490.068
30	4421.536	4363.184	4376.912	4470.63	4489.763
35	4421.202	4362.813	4376.486	4470.266	4489.543
40	4420.952	4362.534	4376.166	4469.992	4489.378
45	4420.756	4362.317	4375.917	4469.778	4489.25
50	4420.6	4362.142	4375.717	4469.607	4489.147

Dr Mukender Singh KADYAN
Department of Statistics and O.R.,
Kurukshetra University, Kurukshetra
Haryana, India 136119
e-mail: mskadian@kuk.ac.in

Article citation info:
GRABOŚ A, BORYGA M. Trajectory planning of end-effector with intermediate point. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 182–187.

Andrzej GRABOŚ
Marek BORYGA

TRAJECTORY PLANNING OF END-EFFECTOR WITH INTERMEDIATE POINT

PLANOWANIE TRAJEKTORII RUCHU CHWYTAKA Z PUNKTEM POŚREDNIM*

The article presents the Polynomial Cross Method (PCM) for trajectory planning of an end-effector with an intermediate point. The PCM is applicable for designing robot end-effector motion, whose path is composed of two rectilinear segments. Acceleration profile on both segments was described by the 7^{th}-degree polynomial. The study depicts an algorithm for the method and the research results presented as the runs of resultant velocity, acceleration and linear jerk of the stationary coordinate system.

Keywords: *trajectory planning, polynomial acceleration profile, jerk.*

W pracy zaprezentowano metodę PCM (Polynomial Cross Method) do planowania trajektorii ruchu chwytaka z punktem pośrednim. PCM ma zastosowanie do planowania ruchu chwytaka, którego tor składa się z dwóch odcinków prostoliniowych. Profil przyspieszenia na obu odcinkach opisany został wielomianem siódmego stopnia. W pracy przedstawiono algorytm metody oraz wyniki w postaci przebiegów prędkości, przyspieszenia i udaru liniowego.

Słowa kluczowe: *planowanie trajektorii, chwytak, wielomianowy profil przyspieszenia, udar.*

1. Introduction

Trajectory planning proves to be the first and critical phase in the operation of robotic workstations (such as supporting of machines, painting, welding, sealing, gluing, cutting, assembly, palletization and depalletization). This problem has been an active field of research and consequently vast literature addresses the issue. The authors have applied various techniques for trajectory generation. Some of them considered the minimization of adverse jerk that causes the practical limitation of trajectory mapping errors. The works of Visioli [10] and Dyllong and Visioli [3] highlight the unfavourable jerk effects at the initial and final point of the path for the cubic and third-order trigonometric splines. Interestingly, in some cases, jerk reduction was achieved by the fourth-order trigonometric spline introduction. One of the criteria for optimization of motion path design given by Choi et al. [2], was to keep the jerk within the specified limits. The obtained jerk profiles in the kinematic pairs are discontinuous and step shaped. At the initial and final point of the trajectory, the jerk is different from zero. Red [7] using the S-curves applied the constant (but different from zero) jerk values at the transition period between the constant phases of acceleration and deceleration. The analysis of the link acceleration profiles for Puma 560 manipulator presented by Rubio et al. [8] indicates that negative jerk effect in the kinematic pairs occurs at the initial and final point of the trajectory. That agrees with the observations made by Saramago and Ceccarelli [9] in their study on jerk runs in the kinematic joints. According to Huang et al. [5], the jerk profiles in the kinematic pairs at the both start and end points motion are close to zero. The method proposed by Olabi et al. [6], generates smooth jerk limited pattern constrained by the laws of tool motion and taking into account the joints kinematics constraints. Very interesting research results on the jerk runs in kinematic pairs were reported by Gasparetto and Zanotto [4]. They obtained not only continuous jerk for the applied fifth-order-B-splines, but importantly, its values at the start and end path point were equal zero. The higher degree polynomials to describe acceleration profile were applied by Boryga and Graboś [1]. The authors analyzed the runs of velocity, acceleration and jerk for polynomials of the 5^{th}, 7^{th} and 9^{th} degree. On the basis of the simulation tests performed, they achieved the lowest values of the linear and angular jerks for the 7^{th}-degree polynomials.

The authors proposed the Polynomial Cross Method (PCM) algorithm, which allows the design of trajectory comprising two rectilinear segments in the robot workspace. There were formulated the following assumptions concerning the manipulator end-effector motion:

- acceleration profile on both rectilinear segments depicted with 7^{th}-degree polynomial,
- acceleration profile at the initial and end path points is tangent to the time axis that eliminates adverse jerk effect,
- change of run-up phase into brake one occurs at the intermediate point,
- linear acceleration value for any coordinate does not exceed the preset maximum value a_{max},
- end-effector motion proceeds so that resultant velocity does not change at the intermediate point (where rectilinear segments connect).

As a consequence of the presumed constant resultant velocity value at the intermediate point, resultant acceleration is equal to zero. It is noteworthy that at the intermediate point, a direction of the resultant velocity vector gets changed due to the preset path of the end-effector. Substantial advantage of the presented algorithm proves to be a fact that coefficients of the polynomials depicting the acceleration profile on any coordinate are established solely on coordinate increment and preset maximum acceleration. In general, the jerk elimination at the initial and final trajectory point influences the accuracy of trajectory mapping. That appears to be very helpful as far as technological processes such as pick and place, painting, assembly, welding, sealing, gluing, palletization and depalletization are concerned. Layout of the paper comprises the following sections: Section 2 depicting a trajectory planning technique with the 7^{th}-degree polynomial application utilizing the root of an equation multiplicity; Section 3 presenting an algorithm, which was divided into initial computations, computations for a longer and shorter rectilinear segment and final computations; Section 4 demonstrates the example of the proposed algorithm practi-

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

cal employment, while Section 5 summarizes the simulation results. The conclusions are presented in the last section of the paper.

2. Trajectory planning with polynomial use

The planning of robot end-effector trajectory can be accomplished by using higher-degree polynomials, that facilitate acceleration profile development. The study of Boryga and Graboś [1] showed that among the polynomials of 5th, 7th and 9th-degree describing the acceleration profile, the lowest values of linear and angular jerks were reported for the 7th-degree polynomial. Therefore in this paper, the acceleration profile on coordinate x_i is described with the 7th-degree polynomial in the form of

$$\ddot{x}_i(t) = -a_i \cdot (t)^2 \cdot (t - 0.5t_e)^3 \cdot (t - t_e)^2 \tag{1}$$

where: a_i – coefficient of polynomial on coordinate x_i,
$i = 1, 2, 3$ – coordinate number,
t_e – time of motion end.

Acceleration profile described with the 7th-degree polynomial is presented in Fig. 1.

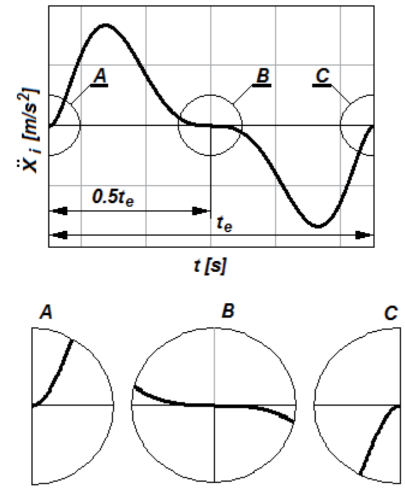


Fig. 1. Acceleration profile described by a 7th - degree polynomial

Acceleration profile is depicted by a continuous function for each coordinate of the Cartesian coordinate system – x_i. Change of the run-up phase into brake phase proceeds at t=0.5t_e and the acceleration profile for t=0, t=0.5t_e and t=t_e is tangent to the time axis. Thus, jerk effect is eliminated at these points. The polynomials describing the profiles of velocity and displacement determined through analytical integration of the dependence (1) go as follows:

$$\dot{x}_i(t) = -a_i \cdot \left(\frac{1}{8}t^8 - \frac{1}{2}t_e t^7 + \frac{19}{24}t_e^2 t^6 - \frac{5}{8}t_e^3 t^5 + \frac{1}{4}t_e^4 t^4 - \frac{1}{24}t_e^5 t^3 \right) \tag{2}$$

$$x_i(t) = -a_i \cdot \left(\frac{1}{72}t^9 - \frac{1}{16}t_e t^8 + \frac{19}{168}t_e^2 t^7 - \frac{5}{48}t_e^3 t^6 + \frac{1}{20}t_e^4 t^5 - \frac{1}{96}t_e^5 t^4 \right) \tag{3}$$

The obtained value $x_i(t)$ is a distance tracked by the robot end-effector on the coordinate i. In order to establish the end-effector coordinate at any moment of time, the following points should be taken into account initial coordinate of end-effector on coordinate i – denoted as x_{bi}, and direction of end-effector motion concordant or discordant with the axis versor orientation. The end-effector coordinate on the coordinate i is defined by the equation

$$x_i(t) = x_{bi} \pm a_i \cdot \left(\frac{1}{72}t^9 - \frac{1}{16}t_e t^8 + \frac{19}{168}t_e^2 t^7 - \frac{5}{48}t_e^3 t^6 + \frac{1}{20}t_e^4 t^5 - \frac{1}{96}t_e^5 t^4 \right) \tag{4}$$

If the end-effector motion is concordant with the versor of the axis i, the plus sign should appear in the equation and the minus one if it is discordant.

3. Planning trajectory with intermediate point

3.1. Polynomial Cross Method (PCM)

PCM is employed to generate end-effector trajectory whose path is composed of two connected rectilinear segments *BM* and *ME* (Fig. 2). Implementation of polynomial acceleration profile of the robot end-effector defined with the equation (1) for the preset segments *BM* and *ME* could cause that the end-effector velocity at the intermediate point was equal to zero. Thus, the problem of trajectory motion planning would be simplified to the motion with a stop at the intermediate point *M*.

For that matter, the ancillary points *E′* and *B′* are introduced. The *E′* point arises from the axial symmetry of point *B* reflected across the intermediate point *M*, whereas the point *B′* through the axial symmetry of the point *E* reflected across the intermediate point *M*. Description of acceleration defined by the equation (1) includes the segments *BE′* and *B′E* (termed total segments in the algorithm). On both total segments, change of the run-up phase into brake one proceeds at the intermediate point *M*. Maximum acceleration of robot end-effector was limited to a_{max} value. It was assumed that at the transition from

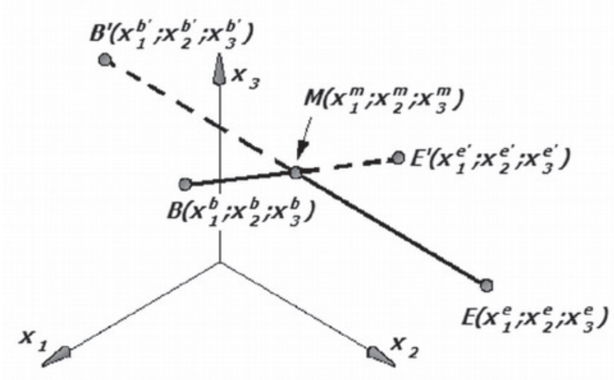


Fig. 2. Planned trajectory BME and ME'and B'M ancillary segments

the *BM* segment to the *ME* one (at the intermediate point *M*), the resultant velocity does not change, while resultant acceleration is equal to zero. Change of the direction and orientation of the velocity vector at the point *M* is imposed by the predetermined trajectory of end-effector motion. At the intermediate point *M*, there occured a rotation of the resultant velocity vector from the *BM* direction towards the *ME* direction. The problem can be solved through the introduction of the arc connecting the rectilinear segments or an alternative stop in the intermediate point. Coefficients of polynomials depicting acceleration profile on the total segments are determined separately for each coordinate x_i. The motion time is calculated using only the path increments and preset maximum acceleration – a_{max}. Velocity value at the intermediate point is established performing the substition of t=0.5t_e into the dependence describing velocity profile (2). The resultant velocity vector at the intermediate point displaces from one segment to the other and projects on the axes of the stationary coordinate system. That facilitates the determination of coefficients of a polynomial depicting aceleration profile on the other total segment. As the motion time on both total segments may vary, it is necessary to perform an appropriate translation in the time of acceleration, velocity, displacement and jerk profile.

3.2. PCM algorithm

3.2.1. Initial computations

Step 1. Assumption of coordinates of the initial, intermediate and final points are denoted by $B(x_1^b;x_2^b;x_3^b)$, $M(x_1^m;x_2^m;x_3^m)$ and $E(x_1^e;x_2^e;x_3^e)$. The points should belong to the workspace.

Step 2. Determination of the coordinates of ancillary points $B'(x_1^{b'};x_2^{b'};x_3^{b'})$ and $E'(x_1^{e'};x_2^{e'};x_3^{e'})$ are made on the grounds of dependence

$$x_i^{b'} = 2x_i^m - x_i^e \quad \text{for} \quad i = 1, 2, 3 \tag{5}$$

$$x_i^{e'} = 2x_i^m - x_i^b \quad \text{for} \quad i = 1, 2, 3 \tag{6}$$

The B' and E' ancillary points need not belong to the workspace. The ancillary distances $B'M$ and ME' are used only to construct an appropriate form of the acceleration profile.

Step 3. Determination of path increments on each coordinate of the total segments

$$\Delta x_i^{BE'} = \left| x_i^{e'} - x_i^b \right| \quad \text{for} \quad i = 1, 2, 3 \tag{7}$$

$$\Delta x_i^{B'E} = \left| x_i^e - x_i^{b'} \right| \quad \text{for} \quad i = 1, 2, 3 \tag{8}$$

Step 4. Scheduling of the coordinate increments starting from the highest, with denotation by subscript in the brackets, in the schedule sequence

$$\Delta x_{\{1\}}^{BE'} \ge \Delta x_{\{2\}}^{BE'} \ge \Delta x_{\{3\}}^{BE'} \tag{9}$$

$$\Delta x_{\{1\}}^{B'E} \ge \Delta x_{\{2\}}^{B'E} \ge \Delta x_{\{3\}}^{B'E} \tag{10}$$

Step 5. Determination of maximum coordinate increment out of BE' and $B'E$ segments and denoting it as $\Delta x_{\{1\}}^L$, that is,

$$\Delta x_{\{1\}}^L = \max\{\Delta x_{\{1\}}^{BE'}, \Delta x_{\{1\}}^{B'E}\} \tag{11}$$

In the $\Delta x_{\{1\}}^L$ denotation, a superscript describes the longer total segment. If $\Delta x_{\{1\}}^{BE'} = \Delta x_{\{1\}}^{B'E}$ increments are equal then $\Delta x_{\{1\}}^{BE'} = \Delta x_{\{1\}}^L$.

Step 6. Assumption of end-effector maximum acceleration a_{max} on the coordinate of the maximum path increment, $\Delta x_{\{1\}}^L$. Thus, the accelerations on the other coordinates will not exceed the preset acceleration a_{max} that results from lower or equal path increments on these coordinates.

3.2.2. Computations longer total segment (L)

Step 1. Determination of polynomial $a_{\{1\}}^L$ coefficient and the end time of motion – t_e^L on coordinate $x_{\{1\}}^L$ requires solution of the equation system

$$\begin{cases} \dfrac{1}{c_1} a_{\{1\}}^L (t_e^L)^9 = \Delta x_{\{1\}}^L \\ -a_{\{1\}}^L (c_2 t_e^L)^2 (c_2 t_e^L - 0.5 t_e^L)^3 (c_2 t_e^L - t_e^L)^2 = a_{\max} \end{cases} \tag{12}$$

Having solved the above equation system, the below was obtained:

$$a_{\{1\}}^L = \frac{c_1}{(\sqrt{c_1 c_3})^9} \cdot \frac{a_{\max}^5}{\sqrt{a_{\max}}\,(\Delta x_{\{1\}}^L)^3 \sqrt{\Delta x_{\{1\}}^L}} \tag{13}$$

$$t_e^L = \frac{\sqrt{c_1 c_3}\sqrt{a_{\max}\Delta x_{\{1\}}^L}}{a_{\max}} \tag{14}$$

where:

$$c_1 = 10080\,,\ c_2 = \frac{1}{2} - \frac{1}{14}\sqrt{21}\,,\ c_3 = -\frac{1}{8} c_2^2 (2c_2 - 1)^3 (c_2 - 1)^2$$

Step 2. Determination of polynomial coefficients for the other coordinates [1]

$$a_{\{i\}}^L = \frac{c_1}{(t_e^L)^9} \Delta x_{\{i\}}^L \quad \text{for} \quad i = 2, 3 \tag{15}$$

Step 3. Determination of components and end-effector resultant velocity at M point

$$\dot{x}_{\{i\}M}^L = \frac{(t_e^L)^8}{6144} a_{\{i\}}^L \quad \text{for} \quad i = 1, 2, 3 \tag{16}$$

$$\dot{x}_M^L = \frac{(t_e^L)^8}{6144} \sqrt{\sum_{i=1}^{3} (a_{\{i\}}^L)^2} \tag{17}$$

3.2.3. Computations shorter total segment (S)

Step 1. Determination of direction cosines between the speed vector in the M point and the axes of the stationary coordinate system

$$\cos(\alpha_i) = \frac{x_i^{e'} - x_i^b}{\sqrt{\sum_{i=1}^{3} (x_i^{e'} - x_i^b)^2}} \quad \text{if} \quad \Delta x_{\{1\}}^{B'E} > \Delta x_{\{1\}}^{BE'} \tag{18}$$

$$\cos(\alpha_i) = \frac{x_i^e - x_i^{b'}}{\sqrt{\sum_{i=1}^{3} (x_i^e - x_i^{b'})^2}} \quad \text{if} \quad \Delta x_{\{1\}}^{B'E} \le \Delta x_{\{1\}}^{BE'} \tag{19}$$

where: α_I– angle between the speed vector at M point and axis x_i of the stationary coordinate system.
Speed vector orientation comes from a motion direction on a segment.

Step 2. Determination of speed components in M point

$$\dot{x}_{iM}^{S} = \dot{x}_M \cdot \cos(\alpha_i) \quad \text{for } i = 1, 2, 3 \tag{20}$$

According to the assumption, the resultant speed value in the M point $\dot{x}_M^L = \dot{x}_M^S = \dot{x}_M$, does not change.

Step 3. Determination of motion time for shorter total segment

$$t_e^S = \frac{105}{64}\frac{\Delta x_i^S}{\dot{x}_{iM}^S} \quad \text{for } i = 1 \tag{21}$$

Formula (21) results from a system of equations formed from the dependences (15) and (16). The same motion time is obtained when appropriate coordinate increments and appropriate velocity components at point M are substituted simultaneously.

Step 4. Determination of polynomial coefficients on each coordinate

$$a_i^S = \frac{6144}{(t_e^S)^8}\dot{x}_{iM}^S \qquad \text{for } i = 1, 2, 3 \tag{22}$$

3.2.4. Final computations

Step 1. Determination of motion start time on the $B'E$ segment

$$t_b = \pm\frac{t_e^L - t_e^S}{2} \tag{23}$$

The time is established so as to obtain the same velocity at exactly the same moment in the M point for both move segments. If $\Delta x_{\{1\}}^{B'E} > \Delta x_{\{1\}}^{BE'}$, dependence (23) should acquire the minus sign, the opposite case should acquire the plus sign.

Step 2. Time displacement of the polynomial depicting the acceleration profile on the coordinates of $B'E$ distance by t_b value

$$\ddot{x}_i(t) = -a_i^L \cdot (t - t_b)^2 \cdot (t - 0.5t_e^L - t_b)^3 \cdot (t - t_e^L - t_b)^2 \ \text{ if } \Delta x_{\{1\}}^{B'E} > \Delta x_{\{1\}}^{BE'} \tag{24}$$

$$\ddot{x}_i(t) = -a_i^S \cdot (t - t_b)^2 \cdot (t - 0.5t_e^S - t_b)^3 \cdot (t - t_e^S - t_b)^2 \ \text{ if } \Delta x_{\{1\}}^{B'E} \le \Delta x_{\{1\}}^{BE'} \tag{25}$$

Analogical time displacement should be done for polynomials, describing the level of velocity, displacement, and jerk.

Step 3. Determination of motion time on the segments along the BME path

$$t_e = \frac{t_e^L + t_e^S}{2} \tag{26}$$

Table 1. The point coordinates for the planned trajectory and ancillary points

Point denotation	Point coordinates [m]		
	x_1	x_2	x_3
B	0.5	0.5	1
M	0.5	0.75	1.25
E	0.75	0.75	1.5
B′	0.25	0.75	1
E′	0.5	1	1.5

4. Numerical example

The point coordinates for the planned end-effector trajectory B, M, E and ancillary points E' and B' are presented in Table 1. The path increments on each coordinate are: $\Delta x_1^{BE'} = 0$, $\Delta x_2^{BE'} = \Delta x_3^{BE'} = 0.5\,m$, $\Delta x_1^{B'E} = \Delta x_3^{B'E} = 0.5\,m$, $\Delta x_2^{B'E} = 0$. Since $\Delta x_{\{1\}}^{BE'} = \Delta x_{\{1\}}^{B'E}$, BE' will be the first segment to study.

The maximum acceleration set is $a_{max} = 2\ m/s^2$ on coordinate x_2. The polynomial coefficients depicting acceleration level on each coordinate as well as motion time go as following: $a_{\{1\}}^L = a_{\{2\}}^L = 354.616\,m/s^9$, $a_{\{3\}}^L = 0$, $t_e^L = 1.343\,s$. As for the BE' distance, path increments are recorded for the coordinates x_2 and x_3 so consequently, the established coefficients $a_{\{1\}}^L$ and $a_{\{2\}}^L$ refer to these coordinates. Resultant speed in M point is $\dot{x}_M^L = 0.864\,m/s$. The direction cosines of a speed vector in M point for the $B'E$ segment go as follows $\cos\alpha_1 = \sqrt{2}/2$, $\cos\alpha_2 = 0$, $\cos\alpha_3 = \sqrt{2}/2$. The velocity components in M point on the $B'E$ segment are $\dot{x}_{1M}^S = \dot{x}_{3M}^S = 0.611\,m/s$, $\dot{x}_{2M}^S = 0$, whereas 1, 2 and 3 indices refer to the axis of the stationary coordinate system. Polynomial coefficients describing acceleration profile on each coordinate of the $B'E$ segment are $a_1^S = a_3^S = 354.616\,m/s^9$, $a_2^S = 0$. Move time recorded on the $B'E$ segment was $t_e^S = 1.343\,s$. Time of motion along the BME path is $t_e = 1.343\,s$, while $t_b = 0$.

5. Simulation tests results

According to the simulation tests performed, the following courses of kinematic characteristics of end-effector were recorded (Fig. 3 – 5). In each figure presented, a continuous line indicates the runs of kinematic characteristics of motion for the designed trajectory BME, while a dashed line – the courses for ancillary segments (ME' and $B'M$). A planned motion path and ancillary distances are displayed in the stationary coordinate system $x_1x_2x_3$, whereas kinematic characteristics of motion at two planes perpendicular to the plane determined by the points of the generated BME trajectory.

Fig. 3 displays the runs of end-effector speeds on the BE' and $B'E$ segments. The maximum velocity $\dot{x}_M^L = \dot{x}_M^S = 0.864\,m/s$ is obtained at the point M. The speed value at transition from the BM segment to ME did not change, while a direction of the resultant velocity vec-

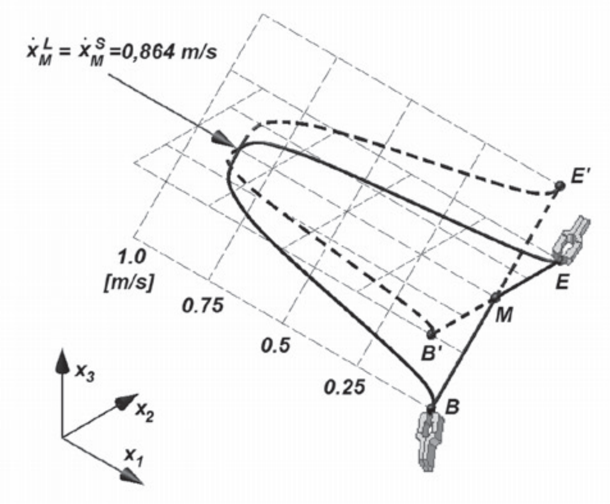


Fig. 3. Resultant velocity course along planned BME path and ancillary segments

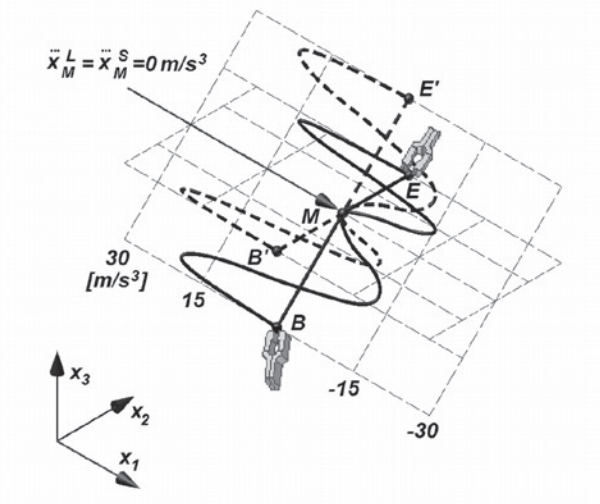


Fig. 5. Resultant jerk course along planned BME path and ancillary segments

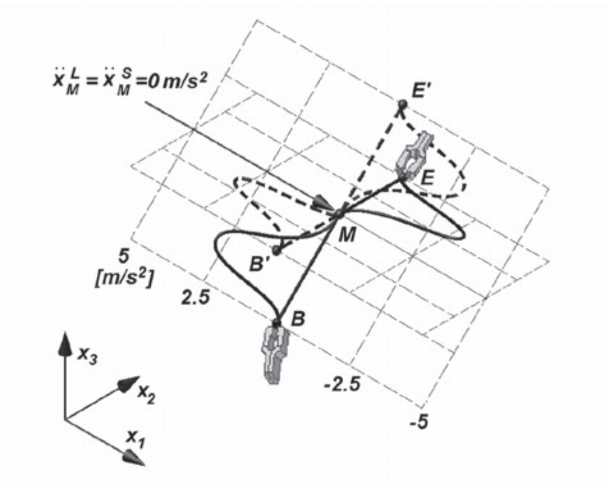


Fig. 4. Resultant acceleration course along planned BME trajectory and ancillary segments

tor changes from *BM* path to *ME*. The runs of end-effector resultant linear acceleration on the distances *BE'* and *B'E* are presented in Fig. 4. In the points *B*, *M*, *E* acceleration is equal to zero. The obtained absolute maximum acceleration at each coordinate does not surpass the set value $a_{max} = 2\ m/s^2$ and the maximum resultant value reaches 2.83 m/s^2. A jerk value at points *B* and *E* is equal zero (Fig. 5). The maximum jerk value is 19.8 m/s^3.

6. Conclusions

On the basis of the simulation tests of the manipulator end-effector motion according to the PCM, the following conclusions were formulated:

a) The profiles of resultant velocity, acceleration and jerk obtained by the PCM application are continuous on the *BM* and *ME* segments. At the intermediate point *M*, resultant velocity value does not change, consistently with the underlying assumption. The velocity vector direction changes according to the motion direction (from *BM* towards *ME*).
b) Generation of trajectory according to the PCM may be utilized in some technological processes (pick and place, painting, assembly, welding, sealing, gluing, palletization and depalletization), where it is critical to eliminate jerk effect in the initial and final point of the trajectory.
c) If deformability of kinematic chain occurs, jerk elimination will result in vibration limitation that guarantees lower tracking errors.

Our further research will focus on the effect of trajectory of the robot manipulator end-effector (planned using PCM) on kinematics and manipulator dynamics.

References

1. Boryga M, Graboś A. Planning of manipulator motion trajectory with higher-degree polynomials use. Mechanism and Machine Theory 2009; 44: 1400–1419.
2. Choi YK, Park JH, Kim HS, Kim JH. Optimal trajectory planning and sliding mode control for robots using evolution strategy. Robotica 2000; 18: 423–428.
3. Dyllong E, Visioli A. Planning and real-time modifications of a trajectory using spline techniques. Robotica 2003; 1: 475–482.
4. Gasparetto A, Zanotto V. Optimal trajectory planning for industrial robots. Advances in Engineering Software 2010; 41: 548–556.
5. Huang P, Xu Y, Liang B. Global minimum-jerk trajectory planning of space manipulator. International Journal of Control, Automation, and Systems 2006; 4 no. 4: 405–413.
6. Olabi A, Béarée R, Gibaru O, Damak M. Feedrate planning for machining with industrial six-axis robots. Control Engineering Practice 2010; 18(5): 471–482.
7. Red E. A dynamic optimal trajectory generator for Cartesian Path following. Robotica 2000; 18: 451–458.
8. Rubio FJ, Valero FJ, Suñer JL, Mata V. Simultaneous algorithm to solve the trajectory planning problem. Mechanism and Machine Theory 2009; 44: 1910–1922.

9. Saramago SFP, Ceccarelli M. An optimum robot path planning with payload constraints. Robotica 2002; 20: 395–404.
10. Visioli A. Trajectory planning of robot manipulators by using algebraic and trigonometric splines. Robotica 2000; 18: 611–631.

Andrzej GRABOŚ, Ph.D. (Eng.)
Marek BORYGA, Ph.D. (Eng.)
The Department of Mechanical Engineering and Automation,
The Faculty of Production Engineering, University of Life Sciences in Lublin,
Doświadczalna str. 50A, 20-280 Lublin, Poland
e-mails: andrzej.grabos@up.lublin.pl, marek.boryga@up.lublin.pl

Article citation info:
GUO Y-M, RAN C-B, LI X-L, MA J-Z, ZHANG L. Weighted prediction method with multiple time series using multi-kernel least squares support vector regression. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 188–194.

Yang-Ming GUO
Cong-Bao RAN
Xiao-Lei LI
Jie-Zhong MA
Lu ZHANG

WEIGHTED PREDICTION METHOD WITH MULTIPLE TIME SERIES USING MULTI-KERNEL LEAST SQUARES SUPPORT VECTOR REGRESSION

METODA WAŻONEJ PREDYKCJI WIELOKROTNYCH SZEREGÓW CZASOWYCH Z WYKORZYSTANIEM WIELOJĄDROWEJ REGRESJI WEKTORÓW WSPIERAJĄCYCH METODĄ NAJMNIEJSZYCH KWADRATÓW (LS-SVR)

Least squares support vector regression (LS-SVR) has been widely applied in time series prediction. Based on the case that one fault mode may be represented by multiple relevant time series, we utilize multiple time series to enrich the prediction information hiding in time series data, and use multi-kernel to fully map the information into high dimensional feature space, then a weighted time series prediction method with multi-kernel LS-SVR is proposed to attain better prediction performance in this paper. The main contributions of this method include three parts. Firstly, a simple approach is proposed to determine the combining weights of multiple basis kernels; Secondly, the internal correlative levels of multiple relevant time series are computed to present the different contributions of prediction results; Thirdly, we propose a new weight function to describe each data's different effect on the prediction accuracy. The experiment results indicate the effectiveness of the proposed method in both better prediction accuracy and less computation time. It maybe has more application value.

Keywords: *time series, weighted prediction, least squares support vector regression (LS-SVR), multiple kernel learning (MKL).*

Regresja wektorów wspierających metodą najmniejszych kwadratów (LS-SVR) jest szeroko stosowana w predykcji szeregów czasowych. Opierając się na fakcie, że jeden rodzaj niezdatności może być reprezentowany przez wiele relewantnych szeregów czasowych, w niniejszej pracy wykorzystano wielokrotne szeregi czasowe do wzbogacenia informacji predykcyjnych ukrytych w szeregach czasowych oraz posłużono się metodą uczenia wielojądrowego (multi-kernel) w celu mapowania informacji do wysoko wymiarowej przestrzeni cech, a następnie zaproponowano metodę ważonej predykcji wielokrotnych szeregów czasowych z wykorzystaniem wielojądrowej regresji LS-SVR służącą osiągnięciu lepszej wydajności prognozowania.Metoda składa się z trzech głównych części. Po pierwsze, zaproponowano prosty sposób określania łącznej wagi wielu jąder podstawowych. Po drugie, obliczono wewnętrzne poziomy korelacyjne wielokrotnych szeregów czasowych w celu przedstawienia różnego udziału wyników prognozowania. Po trzecie, zaproponowano nową funkcję wagi do opisu różnego wpływu poszczególnych danych na trafność predykcji. Wyniki doświadczenia wskazują na skuteczność proponowanej metody zarówno jeśli chodzi o lepszą trafność predykcji jak i krótszy czas obliczeniowy. Proponowane rozwiązanie ma potencjalnie dużą wartość aplikacyjną.

Słowa kluczowe: *szereg czasowy, predykcja ważona, regresja wektorów wspierających metodą najmniejszych kwadratów (LS-SVR), uczenie wielojądrowe (MKL).*

1. Introduction

Fault or health trend prediction technique has become one of the effective ways to protect the safe operation of high reliable systems. However complex systems often show complex dynamic behaviors and uncertainty, which lead to hardly establishing their precise physical models. In this case, in order to obtain the satisfactory prediction results, time series analysis methods are often used to perform the prediction in practice [2, 12, 15, 19, 26]. Among the known non-linear time series prediction methods, the effectiveness of statistics theory based methods have been demonstrated, such as Artificial Neural Networks (ANN), Support Vector Regression (SVR), etc.

ANN has been applied in many fields due to its universal approximation property. However ANN suffers from local minimum traps, difficulty in determining the hidden layer size and learning rate, poor capacity for generalization, etc. [8, 10, 32] On the contrary, SVR overcomes the problems existing in ANN. SVR aims at the global optimum and exhibits better accuracy in non-linear and non-stationary time series data prediction due to its implementation of the structural risk minimization principle [10, 27, 28]. But complexity of SVR depends not only on the input space dimension, but also on the number of sample data. For large sample data, the quadratic programming (QP) problem is more complex, it will cost a lot of computing time. For this reason, LS-SVR was proposed by Suykens et al. [16, 23] In

LS-SVR, the inequality constrains are replaced by equality constrains. This way, solving a QP is converted into solving linear equations, and the calculation time is reduced significantly. Thus, LS-SVR attracts more attention in time series prediction [5, 6, 19, 20, 26, 33].

In many applications of fault or health condition prediction, one certain condition may be represented by one major variable and several relevant variables. In order to achieve satisfactory prediction, these auxiliary time series relating to the major time series are utilized to enrich the information and improve the prediction accuracy. In this case, how to fully present the information hiding in the multiple time series data becomes a key issue. The kernel function is used to map the input data to high dimensional feature space, so it influences the learning performance of LS-SVR, that means a appropriate kernel function can more fully present the information in time series data. However, LS-SVR with a single kernel function is not a good choice to all the data sets, especially for multiple time series data, although the kernel parameters can be optimally chosen to enhance the generalization capability.

Some researchers applied Multiple Kernel Learning (MKL) to solve the above problems [13, 29]. MKL provides a more flexible framework than single kernel. Under the framework, the information in time series data can be mined more adaptively and effectively, i.e., MKL explicitly learns the weights of basis kernels from different time series data sources, and the relationships among them are learned meanwhile. Moreover, MKL can avoid the difficulty of appropriate kernel function selection. Thus, multi-kernel LS-SVR has better prediction accuracy in practice [13,14].

However in order to obtain better prediction results, some problems which accord to the requirements of applications, should still be considered, such as fast and accurate prediction.

(1) In MKL framework, the time series data samples are generally learned by a linear convex combination of basis kernels. The reported methods of determining the combining weights of basis kernels, such as software packages [1] and joint optimization selection algorithm [9, 34], are always complex. They are generally unapt for applications.

(2) Although some researchers also used multiple relevant time series to perform prediction [6, 17, 31, 35, 36], different interrelated levels between major time series and auxiliary time series have different influences on prediction accuracy. It is necessary to determine the interrelated levels between them, and they represent the weight values of each time series for the prediction.

(3) The original prediction methods always assume that all the training time series data have same contribution to the prediction. According to the new information principle [3], the data near the current prediction point will affect the prediction much more. Thus, in order to achieve more accurate results, each sample data should be weighted according to their distance far from the current prediction point.

Thus a weighted prediction method with multi-kernel LS-SVR using multiple relevant time series is proposed in this paper. According to the application requirements, we apply three ways to achieve better prediction results. One is to compute correlative levels of multiple relevant time series to represent their different contributions to prediction results; Secondly, we propose a weight function to present the different influence of each history data on prediction; Finally, we establish a new multi-kernel LS-SVR based on time-distance-weighted factor of each time series, and in order to improve the application value of the proposed method, a simple approach of determining the combining weights of the multiple basis kernels is proposed to reduce the calculation time.

The rest of the paper is organized as follows: Section 2 gives a brief review of LS-VR and multiple kernel learning (MKL) algorithm; Section 3 proposes the weighted prediction method which includes three computational approaches: (1) combination coefficients of multiple basis kernels, (2) correlative levels of the multiple time series, and (3) time-distance-weighted factors of each time series data; Section 4 shows simulation and application experiments; and the conclusions are drawn in Section 5.

2. A brief review of related work

2.1. Least squares support vector regression

LS-SVR has many advantages, such as simpler algorithm, faster operation speed, etc. It is widely applied in regression. The goal of LS-SVR is to estimate a function that is as "close" as possible to the target values for every data point, and at same time, is as "flat" as possible for good generalization. The regression principle of LS-SVR can be expressed as follows.

Consider a training data set of n data points $\{x_i, y_i\}_{i=1}^n$ with input data $x_i \in R^d$, and $y_i \in R$ is the corresponding output or target value. LS-SVR is to construct the regression function with the following form

$$f(x) = y = w^T \varphi(x) + b \tag{1}$$

where $\varphi(\cdot)$ is used to non-linearly map the input data to the high dimensional feature space , w is the weight vector and b is the bias term.

According to the structural risk minimization principle[27,28], the function regression problem can be represented as a constraint optimization problem as follows

$$\begin{aligned} &\min \quad J(w,b) = \frac{1}{2}\|w\|^2 + \frac{c}{2}\sum_{i=1}^{n} e_i^2 \\ &s.t. \quad y_i = w^T \varphi(x_i) + b + e_i \end{aligned} \tag{2}$$

where $i = 1,2,...,n$, $J(w,b)$ is the cost function, c is a positive real constant (regularization parameter) and $e_i \in R$ is an error variable.

In order to solve the above constraint optimization problem, the Lagrangian function is constructed by transforming constraint optimization problems into unconstraint ones

$$L(w,b,e,\alpha) = \frac{1}{2}w^T w + \frac{1}{2}c\sum_{i=1}^{n} e_i^2 - \sum_{i=1}^{n} \alpha_i (w^T \varphi(x) + b + e_i - y_i) \tag{3}$$

where α_i is the i-th Lagrange multiplier. It is obvious that the optimal solution of Eq.(2) satisfies the Karush-Kuhn-Tucker (KKT) conditions. The optimal conditions are expressed as follows

$$\begin{cases} \dfrac{\partial L}{\partial w} = w - \sum_{i=1}^{n} \alpha_i \varphi(x_i) = 0 \Rightarrow w = \sum_{i=1}^{n} \alpha_i \varphi(x_i) \\ \dfrac{\partial L}{\partial b} = -\sum_{i=1}^{n} \alpha_i = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i = 0 \\ \dfrac{\partial L}{\partial \alpha_i} = w^T \varphi(x_i) + b + e_i - y_i = 0 \Rightarrow y_i = w^T \varphi(x_i) + b + e_i \\ \dfrac{\partial L}{\partial e_i} = ce_i - \alpha_i = 0 \Rightarrow e_i = \dfrac{1}{c}\alpha_i \end{cases} \tag{4}$$

After eliminating w and e_i from Eq.(4), we could obtain the solution by the following linear equations

$$\begin{bmatrix} 0 & \mathbf{1}_n^T \\ \mathbf{1}_n & \mathbf{K}+\mathbf{I}/\mathbf{c} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \tag{5}$$

where $\mathbf{K}(i,j) = k(x_i, x_j) = \varphi(x_i)^T \varphi(x_j)$, $\alpha = [\alpha_1, \alpha_2, ..., \alpha_n]^T$, $\mathbf{1}_n$ is an n-dimensional vector of all ones, $\mathbf{I}$ is a unite matrix and $\mathbf{y} = [y_1, y_2, ..., y_n]^T$. Eq.(5) can be factorized into a positive definite system[18].

Let $\mathbf{H} = \mathbf{K} + \mathbf{I}/c$, we get the equations from Eq.(5)

$$\begin{cases} \mathbf{1}_n^T \alpha = 0 \\ \mathbf{1}_n b + \mathbf{H}\alpha = \mathbf{y} \end{cases} \tag{6}$$

Then Lagrange dual variables α and bias term b are obtained solely by

$$\begin{cases} \alpha = \mathbf{H}^{-1}(\mathbf{y} - \mathbf{1}_n b) \\ b = \mathbf{1}_n^T \mathbf{H}^{-1} \mathbf{y} (\mathbf{1}_n^T \mathbf{H}^{-1} \mathbf{1}_n)^{-1} \end{cases} \tag{7}$$

Any unlabeled input x can be subsequently regression estimation by the following function

$$\hat{y}(x) = \sum_{i=1}^{n} \alpha_i k(x_i, x) + b \tag{8}$$

2.2. Multiple kernel learning algorithm

The selection of kernel function and its corresponding parameters is the key issue for prediction accuracy. However no rules have been reported to guide the selection in theory. In this case, MKL was proposed by Lanckriet, et al. [13]

In MKL framework, a combined kernel function is defined as the weight sum of several individual basis kernels. Researchers proposed a variety of methods to integrate multiple basis kernels [30]. The linear convex combination of basis kernels is most frequently used. In this paper, using the equations described by Sonnenburg et al. [24] We consider the following form of combined kernel

$$\mathbf{K} = \sum_{j=1}^{m} \mu_j \mathbf{K}_j \tag{9}$$

where $\sum_{j=1}^{m} \mu_j = 1$, $\mu_j \geq 0$ ($j = 1, 2, ..., m$), m is the number of basis kernels, μ_j is the combining weight of the j-th basis kernel. Obviously $\mathbf{K}$ is Symmetric Positive Semidefinite Matrix [22], i.e., $\mathbf{K} \succeq 0$. Afterward, all kernel matrices $\mathbf{K}_j$ are normalized by replacing $\mathbf{K}_j(x_p, x_q) \Big/ \sqrt{\mathbf{K}_j(x_p, x_p)\mathbf{K}_j(x_q, x_q)}$ to get unit diagonal matrices.

The key of MKL is to obtain the optimal combining weights μ_j. This problem can be solved as a QCQP problem[34] efficiently by general-purpose optimization software packages[14]. Moreover, some researchers also applied joint optimization selection algorithms to obtain the combining weights μ_j and the parameters of LS-SVR simultaneously. But all the solution methods are complex in practice. Thus, we proposed a simple method to fix this problem in this paper.

3. Proposed weighted multiple time series prediction method

In this section, we propose a new scheme to obtain better prediction performance. Firstly, we use multiple kernel functions consisting of several basis kernels to show the information more effectively in the high dimensional mapping feature space. A simple approximate approach is presented to compute the combining weights with less calculation complexity. Then we propose the weighted prediction method. In the method, we calculate the correlation coefficient of each time series as weight factor, which present the influence factor on prediction accuracy with each time series, and based on the distance of each time series data far from the current prediction point, we weight the effects of the history data on prediction via a modified weight function.

3.1. Combination coefficients of multiple basis kernels

In this paper, we apply the new kernel with a linear combination of basis kernels, shown as Eq.(9). In order to reduce the computing complexity, we propose a simple method to determine combining weights, i.e., the combining weights of basis kernels are determined according to the root mean squared error (RMSE) of each LS-SVR with each single basis kernel. This way, smaller RMSE value will get bigger weight value. The RMSE of multiple time series prediction is defined as follows

$$\sigma_{RMSE} = \sqrt{\frac{1}{MN} \sum_{i=1}^{N} \sum_{k=1}^{M} (y_i(k) - \hat{y}_i(k))^2} \tag{10}$$

where M is the number of the relevant parameters, N is the number of original training sample data, and $y_i(k)$ and $\hat{y}_i(k)$ are the prediction value and actual value respectively. The linear combining weights μ_j can be computed as follows

$$\mu_j = \frac{\sum_{r=1}^{m} \sigma_r - \sigma_j}{(m-1)\sum_{r=1}^{m} \sigma_r} \tag{11}$$

where σ_j is the prediction RMSE of the j-th kernel, $\sum_{r=1}^{m} \sigma_r$ is the sum RMSE of all basis kernels, and $\sum_{r=1}^{m} \sigma_r - \sigma_j$ presents the contribution of the j-th kernel.

Obviously, the proposed method of calculating combination coefficients has less complexity comparing with the methods described in section 2.

3.2. Weight factors of multiple time series

3.2.1. Weight factors of major and auxiliary time series

Multiple relevant time series are used to enrich the information in data. However, each time series has different effects on prediction because they have different degrees of information which represent the system's fault or health condition. In this paper, we select the time series, which mainly represents the system fault or health state, as major time series, the others are auxiliary time series. Then the correlation coefficients between the major time series and the auxiliary time series will be computed, and they will be utilized to improve the prediction accuracy.

The purpose of correlation analysis is to measure and interpret the strength of linear or non-linear relationship between two continuous variables [11, 22]. We select the commonly used correlation coefficients, Pearson correlation coefficient [4, 7], to assess the strength of the relationships of multiple relevant time series. The Pearson correlation coefficient computing formula is shown as follows

$$R = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{12}$$

where R is the correlation coefficient between bivariate data x_i and y_i values ($i = 1,2,...,n$), $\overline{x}$ and $\overline{y}$ are the mean values of the x_i and y_i respectively. The Pearson correlation coefficient may be computed by means of a computer-based statistics program "Microsoft Excel" using the option "Correlation" under the option "Data Analysis Tools". Moreover it can also be calculated by Matlab.

3.2.2. Time-distance-weighted factors of each time series

The time series data closing to the current prediction point have greater relevance to current prediction, on the contrary, less relevance with data far from the current prediction point. Hence, we propose a modified weight function to present the different weight factor of each historical data.

According to Ref.[25] and Ref.[37], consider the generation sample set from the raw time series $\{x_k, y_k\}$ ($k = 1,2,...,n$), we define a new weight function of x_k as follows

$$d_k = e^{-(n-k)^2 / 2\lambda^2}, \quad k = 1,2,...,n \tag{13}$$

where λ is a given parameter, and a small d_i can reduce the storage of historical data and speed up the training. The objective function is expressed as follows

$$\begin{aligned} \min \quad & J(w,e) = \frac{1}{2}w^T w + \frac{1}{2}cd_k \sum_{i=1}^{n} e_k{}^2 \\ s.t. \quad & y_k = w^T \varphi(x_k) + b + e_k, \quad k = 1,2,...,n \end{aligned} \tag{14}$$

Then the Lagrangian function is established below

$$L(w,b,e,\alpha) = \frac{1}{2}w^T w + \frac{1}{2}cd_k \sum_{k=1}^{n} e_k^2 - \sum_{k=1}^{n} \alpha_k (w^T \varphi(x) + b + e_k - y_k) \tag{15}$$

where $\alpha_k \geq 0$ ($k = 1,2,...,n$) are the Lagrangian multipliers.
According to KKT conditions, we can get the following equations

$$\begin{cases} \dfrac{\partial L}{\partial w} = w - \sum_{k=1}^{n} \alpha_k \varphi(x_k) = 0 \Rightarrow w = \sum_{k=1}^{n} \alpha_k \varphi(x_k) \\ \dfrac{\partial L}{\partial b} = -\sum_{k=1}^{n} \alpha_k = 0 \Rightarrow \sum_{k=1}^{n} \alpha_k = 0 \\ \dfrac{\partial L}{\partial \alpha_k} = w^T \varphi(x_k) + b + e_k - y_k = 0 \Rightarrow y_k = w^T \varphi(x_k) + b + e_k \\ \dfrac{\partial L}{\partial e_k} = cd_k e_k - \alpha_k = 0 \Rightarrow e_k = \dfrac{1}{cd_k}\alpha_k \end{cases} \tag{16}$$

And then rewrite Eq.(5) with a new form as follows

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & k(x_1,x_1) + 1/cd_1 & \cdots & k(x_1,x_n) \\ \vdots & \vdots & \cdots & \vdots \\ 1 & k(x_n,x_1) & \cdots & k(x_n,x_n) + 1/cd_n \end{bmatrix} \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix} \tag{17}$$

4. Experiments and results analysis

We conduct two simulation experiments and one application experiment to evaluate the performance of proposed method. The prediction experiments are run 100 times and the averages results are taken. All the experiments adopt MatlabR2011b with LS-SVMlab1.8 Toolbox (The software and guide book can be downloaded from http://www.esat.kuleuven.be/sista/lssvmlab) under Windows XP operating system.

4.1. Simulation experiments and results analysis

The simulation experiments include Experiment I and Experiment II. They are conducted to test the proposed method presented in section 3. All the simulation experiments are performed using Lorenz function, because Lorenz function is a typical time series and its variables depend on each other. Lorenz function's corresponding differential equations are shown as follows

$$\begin{cases} x' = -ax + yz \\ y' = -b(y - z) \\ z' = -xy + cy - z \end{cases}$$

Let $a = 8/3$, $b = 10$, $c = 28$, range of initialization as [1,1,1], and simulation step as 0.1 with Fourth-oder Runge-Kutta method. We collect 800 data of the three time series of x (major time series), y and z (auxiliary time series) respectively. We select the first 400 data of x, y and z time series as training data and the last 400 time series data are testing data. In addition, we apply C-C method[21] to generate training sample sets because Lorenz time series is chaotic time series. The prediction efficiency depends on the RMSE, training time (TrTime) and prediction time (PrTime). One Gaussian RBF $K(x,y) = \exp(-\frac{\|x-y\|^2}{2\sigma^2})$ and one Linear kernel function $K(x,y) = x^T y$ are adopted as basis kernel functions. All the parameters will be jointly optimized by traditional gridding search method with rang of [0.1, 1000]

In Experiment I, we use variable x time series alone to do prediction with tradition multi-kernel LS-SVR reported in Ref.[35] and Ref.[36]. This experiment compares the following two methods: one is obtains the combining weights via optimization software packages (called Method A); the other one is the proposed simple approximate approach in section 3.1(called Method B). The results are shown in Figure1 and Table 1.

From Figure1 and Table 1, we can see that although the prediction accuracy of the new simple computing method (Method B) is adequately reduced comparing with Method A, it also has a good results. Method B can greatly reduce the total computing time, especially

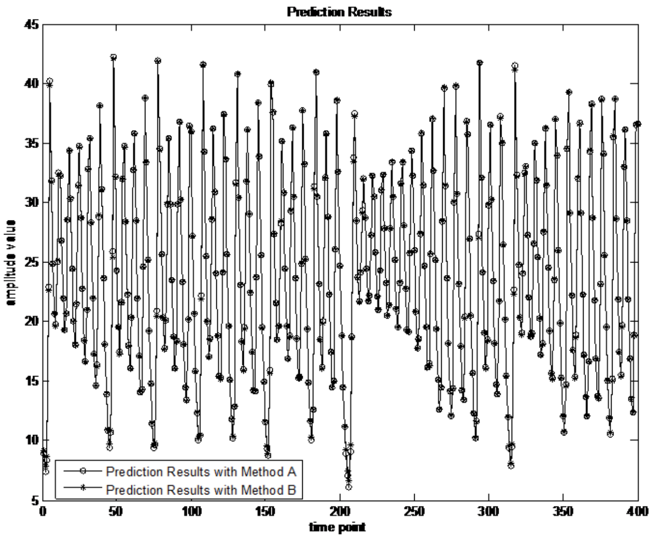


Fig. 1. Prediction Results with Method A and Method B

Table 1. Prediction Results of Method A and Method B

x	TrTime/s	PrTime/s	RMSE
Method A	4.1501	0.2406	2.2258
Method B	3.0480	0.2456	2.8927

training time. The results also indicate that the proposed approximate method is an effective method.

In Experiment II, the variables y and z time series are utilized to enrich the information of variable x time series. We use same multi-kernel LS-SVR model to compare the following methods: Method C doesn't consider the different contributions of each time series and their history data on prediction; Method D applies the proposed approach proposed in section 3.2. Here, we select same kernel functions and optimization method as Experiment I. The results are reported in Table 2, Figure 2 and Table 3.

Figure 2 and Table 3 show that the weighted time series prediction method can improve the prediction accuracy efficiently, and the computing time is not large increase. These are due to that the proposed method takes the different influence factors on prediction accuracy with each auxiliary time series and their history data into account. The other reason is that almost all the middle values at the calculation process of weight factors are already computed and stored in the process of setting up the prediction method.

Table 2. Correlation Coefficient of Time Series

	r_{xy}	r_{xz}
Correlation Coefficient	-0.0581	-0.0348

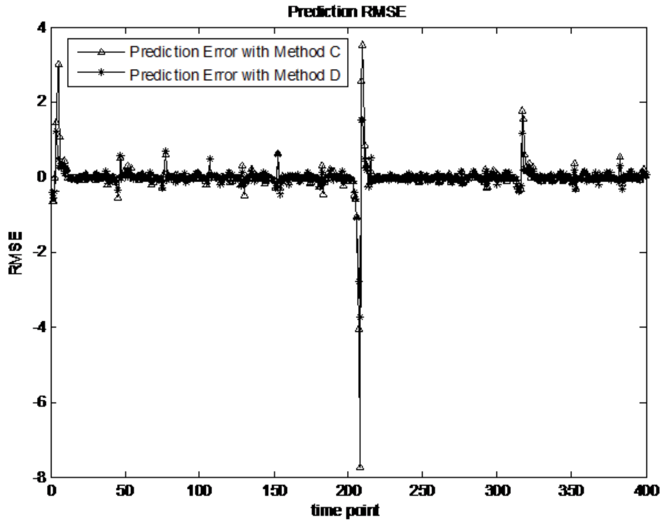


Fig. 2. Prediction Error of Method C and Method D

Table 3. Prediction Results of Method C and Method D

x	TrTime/s	PrTime/s	RMSE
Method C	8.0479	0.8016	2.1765
Method D	9.1241	0.8203	2.0252

4.2. Application experiment and results analysis

We apply the proposed method in a prediction application of one complex avionics system. Four relevant variables time series are collected. They are shown in Figure 3 after preprocessing (omit dimension).

We take the first 15 data of each time series as training samples and look at any continuous 6 as a sample, i.e., the data points from 1 to 15 in the time series are taken as the 10 initial training sample data. The first sample data set consists of points 1 through 6, with the first 5 as the input sample vector and the 6th point as the output. The second sample data set consists of points 2 through 7, with the points 2 through 6 as the input sample vector and the 7th point as the output. This way we have 10 training data out of the first 15 data points.

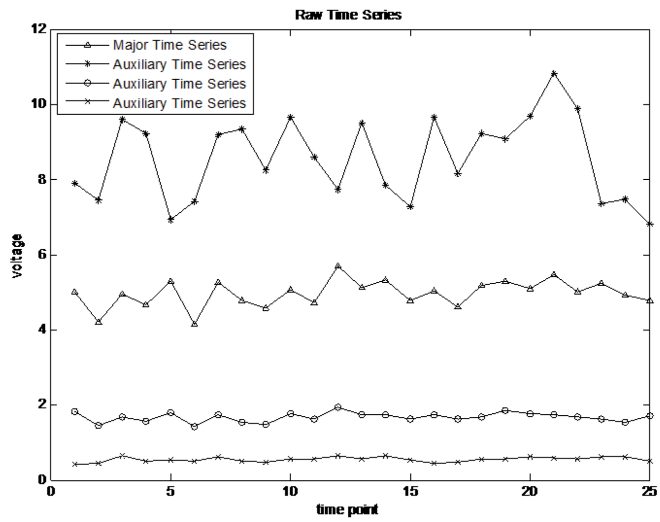


Fig. 3. Raw Time Series of Complex Avionics System

All the parameters are set same as simulation Experiments I and II. The contrast prediction experiment applies Method A and Method E (described in Section 3). The prediction results of the major time series (see Figure 3) are shown in Figure 4 and Figure 5.

In order to show the results clearly, we report them in Table 4 and Table 5.

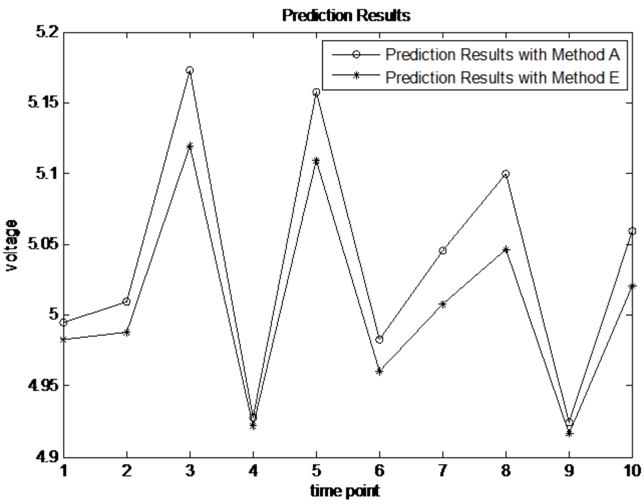


Fig. 4. Results with Method A and Method E

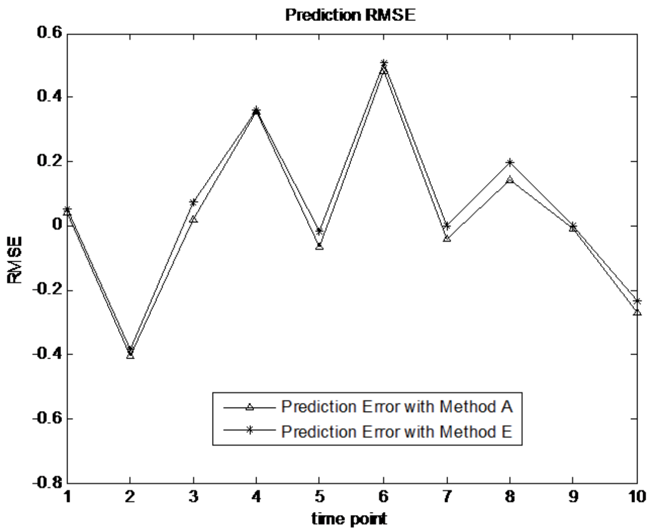


Fig. 5. Error with Method A and Method E

Table 4. Correlation Coefficient of the Time Series

	r_1	r_2	r_3
Correlation Coefficient	0.2791	0.8514	0.6065

Table 5. Prediction Results of Method A and Method E

	TrTime/s	PrTime/s	RMSE
Method A	1.4322	0.0808	0.6771
Method E	1.2927	0.0122	0.5789

From Figure 4, Figure 5 and Table 5, we can see that the proposed method has better prediction results in prediction accuracy and computing time. The results also indicate the proposed method is a good approach, and it can adapt the application better.

5. Conclusions

In this study, we aim at the requirements of applications and analyze the drawbacks of multiple time series prediction by LS-SVR, and then we propose a novel weighted multiple time series prediction method based on multi-kernel LS-SVR. In the new method, we determine the combining weights of each basis kernels by calculating the root mean squared error (RMSE) of prediction using each basis kernel, compute the different contributions to prediction results via correlation analysis between the major time series and auxiliary time series, and make the each historical data with different weight factor based on their distance far from the current prediction point via a modified weight function. The results of simulation and application experiments show that the proposed prediction scheme is an effective approach. It can satisfy the application requirements and may be more valuable in practice.

***Acknowledgments**:*

This work is supported by the National Basic Research Program of China (973 Program), the National Natural Science Foundation of China (No. 61001023 and No. 61101004), Shaanxi Natural Science Foundation (2010JQ8005), Aviation Science Foundation of China (2010ZD53039), and the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

1. Andersen ED, Andersen AD. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. In High performance optimization Norewll, Frenk H, Roos C, Terlaky T, Zhang S (eds.), Kluwer Academic Publishers, 2000.
2. Caesarendra W, Widodo A, Pham Hong Thom, Bo-Suk Yang, Setiawan JD. Combined Probability Approach and Indirect Data-Driven Method for Bearing Degradation Prognostics. IEEE Transactions on Reliability 2011; 60(1):14–20.
3. Deng JL. The primary methods of grey system theory. Wuhan: Huazhong University of Science and Technology Press, 2004.
4. Freund JE. Mathematical statistics (5th ed.) Upper Saddle River, NJ: Prentice Hall,1992.
5. Guo HB, Guan XQ. Application of Least Squares Support Vector Regression in Network Flow Forecasting. The 2nd International Conference on Computer Engineering and Technology, April, 2010.
6. Guo YM, Zhai ZJ, Jiang HM. Weighted prediction of multi-parameter chaotic time series using least squares support vector regression. Journal of Northwestern Polytechnical University 2009; 27(1):83–86.
7. Goldman RN, Weinberg JS. Statistics: an introduction. Upper Saddle River, NJ: Prentice Hall, 1985.

8. Hansen JV, Nelson RD. Neural networks and traditional time series methods: A synergistic combination in state economic forecasts. IEEE transactions on Neural Networks 1997; 8(4):863–873.
9. Jian L, Xia Z H, Liang X J, Gao C H. Design of a multiple kernel learning algorithm for LS-SVM by convex programming. Neural Networks 2011; (24):476–483.
10. Kecman V. Learning and soft computing: support vector machines, neural networks, and fuzzy logic models. Cambridge, MA,USA: MIT Press, 2001.
11. Krzanowski WJ. Principles of multivariate analysis: a user's perspective. Oxford, England: Clarendon, 1988.
12. Liu DT, Wang SJ, Peng Y, Peng XY. Online adaptive status prediction strategy for data-driven fault prognostics of complex systems. 2011 IEEE conference on autotestcon, Sep. 2011.
13. Lanckriet GR, Cristianini N, Bartlett P L, Ghaoui L E, Jordan M I. Learning the kernel matrix with semidefinite programming. Journal of Machine Learning Research 2004; (5):27-72.
14. Li M, Xu JW, Yang JH, Yang DB. Prediction for chaotic time series based on phase reconstruction of multivariate time series. Journal of University of Science and Technology Beijing 2008; 30(2): 208–211,216.
15. Michael P, Rubyca J. A prognostics and health management roadmap for information and electronics-rich systems. IEICE Fundamentals Review 2010; 3(4):25–32.
16. Muller KR, Smola AJ, Ratsch G, et al. Predicting time series with support vector machines. Artificial Neural Networks 1997; 1327(4):999-1004.
17. Nazih AS, Fawwaz E, Osama M A. Medium-term electric load forecasting using multivariable linear and non-Linear regression. Smart Grid and Renewable Energy 2011; 2:126-135.
18. Ojeda F, Suykens J, De MB. Low rank update LS-SVM classifiers for fast variable selection. Neural Network 2008; 21:443-449.
19. Qu J, Zuo MJ. An LSSVR-based algorithm for online system condition prognostics. Expert Systems with Applications 2012; 39(5):6089-6102.
20. Qu J, Zuo MJ. An LSSVR-based machine condition prognostics algorithm for slurry pump systems. Proceedings of the Canadian Society for Mechanical Engineering Forum 2010, June, 2010.
21. Qin Y Q, Cai WD, Yang BR. Research on phase space reconstruction of non-linear time series. Journal of System Simulation 2008; 20(11): 2969–2973.
22. Rodriguez RN, Correlation. In: Kotz S, Johnson N.L., eds. Encyclopedia of statistical sciences. New York, NY: Wiley, 1982.
23. Suykens JAK, Van GT, De BJ, De MB, Vandewaller J. Least squares support vector machines. World Scientific Publishing, 2002.
24. Sonnenburg S, Räosch G, Schäer C, Schökopf B. Large scale multiple kernel learning. Journal of Machine Learning Research 2006; 7:1531–1565.
25. Suykens JAK, Vandewalle J. Least square support vector machines. IEEE Transactions on Circuits and Systems-I 2000; 47(7):1109–1114.
26. Tay FEH, Cao L. Application of support vector machines in financial time series forecasting. The International Journal of Management Science (Omega) 2001; 29: 309–317.
27. Vapnik V. The nature of statistical learning theory. New York, USA Springer Verlag,1995.
28. Vapnik V. Statistical learning theory. John Wiley and Sons, New York, 1998.
29. Wang Z, Chen S, Sun T. MultiK-MHKS: A novel multiple kernel learning algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence 2008;30(2):348–353.
30. Wang H Q, Sun F C, Cai Y N, Chen N, Ding L G. On multiple kernel learning methods. Acta Automatica Sinica 2010; 36(8):1037–1050.
31. Ye M Y, Wang X D, Zhang H R. Chaotic time series forecasting using online least squares support Vector machine regression. Acta Physica Sinica 2005; 54(6):2568-2573.
32. Zhao X H, Wang G, Zhao K K, Tan D J. On-line least squares support vector machine algorithm in gas prdecion. Mining Science and Technology 2009; 19:194-198.
33. Zhao Y H, Zhong P, Wang K N. Application of least squares support vector regression based on time series in prediction of gas. Journal of Convergence Information Technology 2011; 6(1):243-250.
34. Zhang W M, Li C X, Zhong B L. LSSVM parameters optimizing and non-linear system prediction based on cross validation. The fifth International Conference on Natural Computation, Aug, 2009.
35. Zhang X R, Hu L Y, Wang Z S. Multiple kernel support vector regression for economic forecasting. International Conference on Management Science & Engineering (17th), November, 2010.
36. Zhang J F, Hu S S. Chaotic time series prediction based on multi-kernel learning support vector regression. ACTA Physica Sinica (Chinese Physics) 2008; 57(5):2708–2713.
37. Zheng XX, Qian F. Based on the support vector machine online modeling and application. Information and Control 2005; (5):636–640.

Associate Prof. Yang-Ming GUO, Ph.D.
Cong-Bao RAN, Master candidate
Xiao-Lei LI, Master candidate
Associate Prof. Jie-Zhong MA
Lu ZHANG, Master candidate
School of Computer Science and Technology
Northwestern Polytechnical University
Youyi West Road 127, Xi'an Shaanxi, 710072
P. R. China
E-mails: yangming_g@nwpu.edu.cn (Y. M. Guo), fengmingr@gmail.com (C. B. Ran), 465628547@qq.com (X. L. Li), majz@nwpu.edu.cn (J. Z. Ma), 956046544@qq.com (L. Zhang)

Article citation info:
GU H, ZHAO J, ZHANG X. Hybrid methodology of degradation feature extraction for bearing prognostics. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 195–201.

Hongqiang GU
Jianmin ZHAO
Xinghui ZHANG

HYBRID METHODOLOGY OF DEGRADATION FEATURE EXTRACTION FOR BEARING PROGNOSTICS

METODYKA HYBRYDOWA EKSTRAKCJI CECH DEGRADACJI DO ZASTOSOWAŃ W PROGNOZOWANIU CZASU ŻYCIA ŁOŻYSK

Hybrid methodology of degradation feature extraction was presented which may enable prediction of remaining useful life of a product. In this methodology, firstly, the signal was de-noised by wavelet analysis. Then the autoregressive model was used to remove the discrete frequencies from de-noised signal. Further, the residual signal which mainly contained impulsive fault signal was enhanced by minimum entropy deconvolution filter. The kurtosis was extracted which was taken as the feature for prognostics. At last, the empirical mode decomposition was used to reduce fluctuation of feature value and to extract the trend content. A case study was presented to verify the effectiveness of the proposed method.

Keywords: *feature extraction, degradation, signal, bearing.*

Przedstawiono hybrydową metodę ekstrakcji cech degradacji, która umożliwia przewidywanie pozostałego okresu użytkowania produktu. W tej metodyce, sygnał został najpierw odfiltrowany z wykorzystaniem analizy falkowej. Następnie, za pomocą modelu autoregresyjnego usunięto z pozbawionego szumów sygnału częstotliwości dyskretne. W dalszej kolejności, sygnał resztkowy, który zawierał głównie impulsowy sygnał uszkodzenia został wzmocniony z zastosowaniem filtru dekonwolucji minimum entropii. Obliczono kurtozę, którą przyjęto jako cechę w procesie prognozowania. Na koniec, zastosowano empiryczną dekompozycję sygnału (EMD) w celu zmniejszenia wahań wartości cechy oraz w celu ekstrakcji trendu. Studium przypadku demonstruje efektywność proponowanej metody.

Słowa kluczowe: *ekstrakcja cech, degradacja, sygnał, łożysko.*

1. Introduction

Bearing is one of the most critical components in rotating machines and its degradation is one of the most frequent reasons which cause a machine to breakdown. In order to reduce downtime costs caused by bearing failures, condition monitoring may be conducted in which remaining useful lives (RUL) of bearings are predicted and fault bearings will be replaced before occurrence of bearing functional failures.

As for the vibration data of bearings, signal modulation effect and noise are two major challenges in incipient fault detection. The modulation effect can be overcome using envelope analysis. As fault signals are often very weak and masked by noise, de-noising and extracting the useful feature which can reflect the degradation effectively from the weak signal are crucial to RUL prediction. In addition, many vibration signals measured externally on machines are distorted by the transmission paths from the source to the transducer. A method known as minimum entropy deconvolution [10] (MED) can remove the effect of the transmission path and enhance the bearing impulsive fault signal. The MED method was applied to gear diagnostics by Endo and Randall [3], and to bearing diagnostics by Sawalhi [7]. Sometimes, the features extracted from full life cycle data of bearing often fluctuate largely which baffles the RUL prediction. These years, the time series modeling has been used to remove discrete frequency which may cause feature fluctuating. Wang and Wong employed the autoregressive (AR) filter to isolate the impulse-like effect of localized cracks in a gear tooth [8, 9]. Barszcz and Sawalhi applied AR and MED to wind turbines' bearing fault detection [1].

In order to enable the prediction of remaining useful life of bearings, a hybrid methodology of degradation feature extraction was presented in this paper. In the method, wavelet analysis was used for de-noising the original signal. AR and MED are employed for enhancing impulsive fault signal. EMD was used to reduce fluctuation of feature value and to extract the trend content.

2. Methodology

2.1. Procedure of degradation feature extraction method

The procedure of bearing degradation feature extraction is illustrated as Figure 1. There are three phases in this procedure. First, bearing signals are de-noised using Matlab's Wavelet Toolbox which eliminates wavelets whose coefficients are smaller than a certain threshold. Second, discrete frequencies are removed from the de-noised signals after applying AR filter, and the smearing effect of transfer path is also removed by MED. At last, EMD method was applied to extract the trend content from features which fluctuate with large ranges.

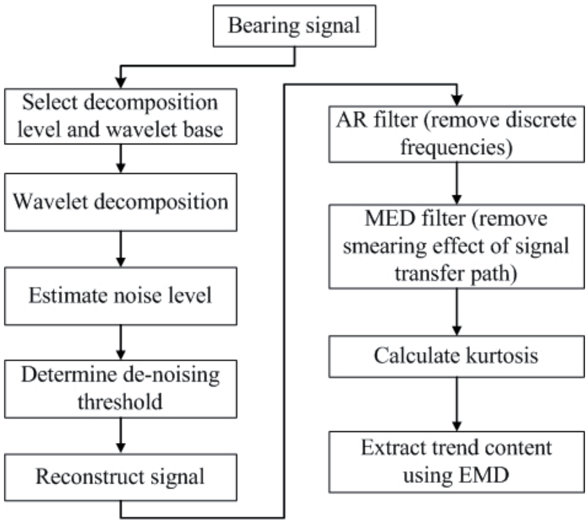


Fig.1. Procedure for degradation feature extraction

2.2. Wavelet based signal de-noising

The basic model for the noisy signal is of the following form:

$$s(n) = f(n) + \sigma\, e(n) \quad n = 0,1\ldots, N-1 \tag{1}$$

Where, $e(n)$ denotes noise. σ denotes the noise level. The objective of de-noising is to suppress the noise part $e(n)$ and to recover $f(n)$. Theoretically, this is implemented by reconstructing the signal from the noisy data such that the mean squared error between $f(n)$ and the reconstructed signal is minimized.

Wavelet de-noising is based on the principle of multi-resolution analysis [2]. By multi-level wavelet decomposition, the discrete detail coefficient and approximation coefficient can be obtained. Grossmann [4] attested that the variance and amplitude of the details of white noise at different levels decreases regularly when the level increases, yet the variance and amplitude of the wavelet transform of the available signal are not related to the change of scale. According to this property, noise can be weakened or even removed by adjusting the wavelet coefficients properly. The de-noising procedure applying in this paper includes following four steps:

- Signal decomposition. Choose a wavelet basis, and choose a level *N*. Compute the wavelet decomposition of the signal at level *N*.
- Estimate the noise level through the detail coefficients of first level.
- Determine the de-noising threshold by penalty strategy.
- Signal reconstruction. Compute wavelet reconstruction using the original approximation coefficients of level *N* and apply soft thresholding to the detail coefficients.

2.3. Fault signal enhancement using AR and MED

One of major sources which mask the relatively weak bearing signals is discrete frequency "noise" from gears or other inherent structures. Even in machines other than gearboxes, there will usually be strong discrete frequency components that may contaminate frequency bands where the bearing signal is dominant. It is generally advantageous therefore to remove such discrete frequency noise before proceeding with bearing fault detection. To this end, AR and MED are employed to enhance the detection of bearings fault, as depicted in Figure 2.

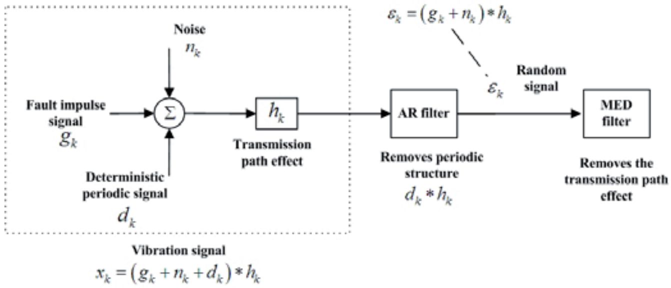


Fig. 2. The filtering process to enhance the detection of bearing faults using AR and MED technique

2.3.1. Removing discrete frequency using AR model

Linear prediction is a basic way which can obtain the model of the deterministic part (vibration of inherent structure of machine) of a signal. In the method, the next value in the series is predicted on the basis of a certain number of previous values. The residual (unpredictable) part of the signal is then obtained by subtraction from the actual signal value. It can be described by the following equation:

$$y(n) = -\sum_{k=1}^{p} a(k)x(n-k) \tag{2}$$

where the predicted current value $y(n)$ is obtained as a weighted sum of the *p* previous values. The actual current value is given by the sum of the predicted value and a noise term:

$$x(n) = y(n) + e(n) \tag{3}$$

The coefficients $a(k)$ can be obtained using the Yule-Walker equations and using Levinson-Durban recursion algorithm.

2.3.2. Enhancing impulsive fault signal using MED

The MED method is designed to reduce the spread of impulse response functions (IRFs), in order to obtain signals closer to the original impulses. It was first proposed by Wiggins to sharpen the reflections from different subterranean layers in seismic analysis. The basic idea of the method is to find an inverse filter which can counteracts the effect of the transmission path, by assuming that the original excitation was impulsive, and then the signal will have high kurtosis. The name of the method derives from the fact that increasing entropy corresponds to increasing disorder, whereas impulsive signals are very structured, requiring all significant frequency components to have zero phase simultaneously at the time of each impulse. Maximizing the structure of the signal corresponds to maximizing the kurtosis of the inverse filter output (corresponding to the original input to the system). According to the Randall's opinion, the method can be also called "maximum kurtosis deconvolution" because the criterion used to optimize the coefficients of the inverse filter is maximization of the kurtosis (impulsiveness) of the inverse filter output. The detailed algorithm can be found in paper of Wiggins [10].

The Figure 2 illustrates the basic idea. The signal $g_k + d_k + n_k$ passes through the structural filter h_k whose output is x_k. The AR filter produces output ε_k, which removes the periodic signal d_k. The inverse (MED) filter produces output y_k, which has to be as close as possible to the fault impulse signal g_k. The MED filter can be modeled as a finite impulsive response filter with L coefficients such that

$$y(k) = \sum_{l=1}^{L} f(l)v(k-l) \tag{4}$$

where f has to invert the system IRFs h such that

$$f * h(k) = \delta\,(k - l_m) \tag{5}$$

The delay l_m is such that the inverse filter can be causal. It will displace the whole signal by l_m but will not change pulse spacing. This method is applied through maximizing the kurtosis of the output signal y_k, by varying the coefficients of the filter f. The kurtosis is taken as the normalized fourth order moment given by

$$O(f) = \frac{\sum_{n=0}^{N-1} (y(k))^4}{\left[\sum_{n=0}^{N-1} (y(k))^2\right]^2} \tag{6}$$

And the maximum is obtained by finding the value of f for which the derivative of the objective function is zero, i.e.,

$$\frac{\partial O(f)}{\partial f} = 0 \tag{7}$$

2.4. EMD based trend component extraction

Empirical mode decomposition (EMD) which is wildly used in nonlinear non-stationary data analysis may be used to extract the content which has some uptrend [5]. It decomposes a raw signal into a set of complete and almost orthogonal components called intrinsic mode functions (IMFs). IMFs represent the natural oscillatory modes embedded in the raw signal. They work as the basis functions which are determined by the raw signal rather than by pre-determined functions. When the trend component is complex which may contain non-period, linear, polynomial, or exponential components, traditional trend component extraction method will be not applicable. For this case, however, the EMD can be used for trend component extraction. According to the EMD, the signal $s(n)$ can be decomposed as:

$$s(n) = \sum_{i=1}^{n} c_i + r_n \tag{8}$$

In the equation, $c_i\ (i = 1, 2, \cdots, n)$ is IMFs, r_n is the residual component. If $s(n)$ denotes the signal which contains the mixture trend component, r_n will be the trend component. The performance of EMD method is shown in the following case study. After applied EMD to features of 7 bearings, the features will be de-noised as shown in Figure 7. It can be seen that all the features progress with time and the trends are very obvious. Also, the trend can reflect the early degradation process which is more significant to RUL prediction.

3. Case Study

A case study was conducted in order to show the performance of the proposed method. The study used PHM 2012 challenge data [6], which is real experiment data characterizing the degradation of ball bearings along their whole operational life. In the experiment, two accelerometers were mounted on the external race of bearing. The first one was placed on the vertical axis and the second was placed on the horizontal axis. Because the radial force was applied on the tested bearing along the horizontal direction, the horizontal sensor can acquire the more effective signals. In this paper, we only analyze the signals collected from horizontal sensor.

3.1. Traditional feature extraction

Investigation on bearing fault prognostics has been implemented recent years. A key to the success of using vibration data for bearing lifecycle prognostics is to develop a relationship between bearing damage and the fault features extracted from the original sensor signals. There are many traditional features such as kurtosis, root mean square (RMS), and the peak value.

As shown in the Figures 3–5, the degradations of bearing 2 and bearing 4 are similar; and the bearing 1 and bearing 3 may have similar characteristic of degradation. However, the peak values and RMS values of bearing 5, bearing 6 and bearing 7 do not increase with bearing's working time. This presents a complex situation where traditional features may not be applicable to predict the RUL of bearing. In order to extract the features which progress with time, AR and MED method will be used in this data analysis which can enhance the fault signal of bearing. However, in the case that the signals have not periodical impulsive which can be seen obviously from the wave, the effectiveness of AR and MED will be significantly influenced by noise. This is because the MED will enhance the some impulsive like noise and signals without de-noising process will be still masked by noise. Therefore, de-noising is very important step in the proposed method, which can be demonstrated through analyzing the end file of bearing 1 as illustrated by Figure 6.

3.2. Feature extraction using the proposed method

The proposed method of feature extraction contains four steps: denoising, AR filtering, MED filtering and kurtosis calculation. Using the method, the features of 7 bearings can be obtained as shown in Figure 7. It can be seen from Figure 7 that all the features are non-stationary which fluctuate with large ranges. Therefore, it is very difficult to predict the RUL in this situation. The characteristics of signals commonly exist in many real cases of condition monitoring and RUL prediction. In order to deal with the problem, it is necessary to extract trend components from non-stationary features. There are many methods which can be used for trend component extraction, such as. linear trend component extraction, power function trend component extraction, exponential trend component extraction, period trend component extraction, and mixture trend component extraction.

In this paper, EMD is employed to de-noise for the features of bearings. De-noised features of 7 bearings based on EMD are shown in

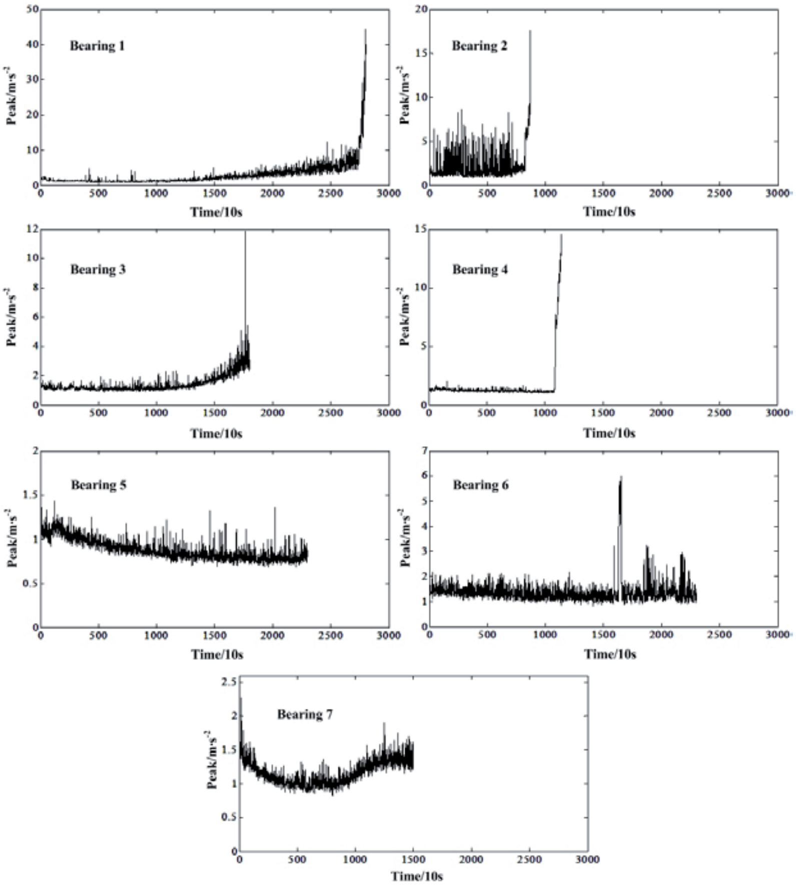


Fig. 3. The peak value of 7 bearings' life vibration under condition 1

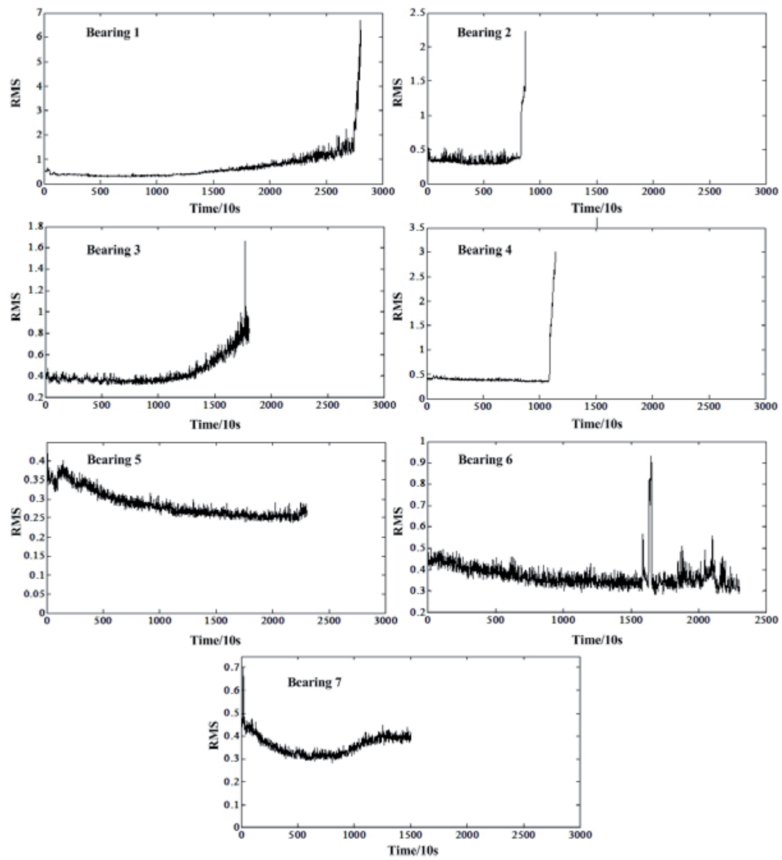


Fig. 4. The RMS value of bearings' vibration under condition 1

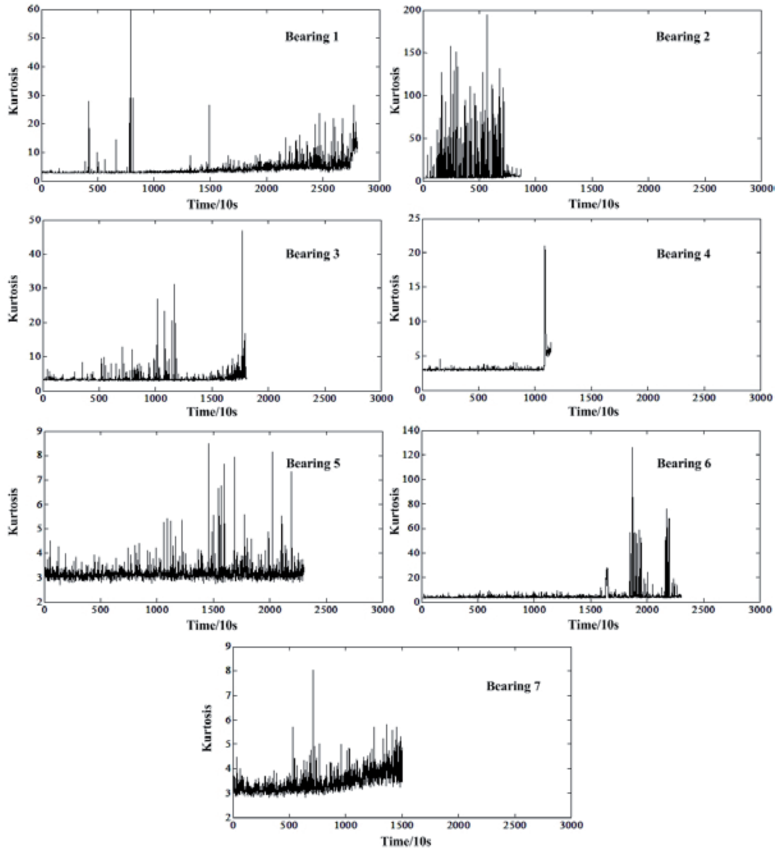


Fig. 5. The Kurtosis value of 7 bearings' vibration under condition 1

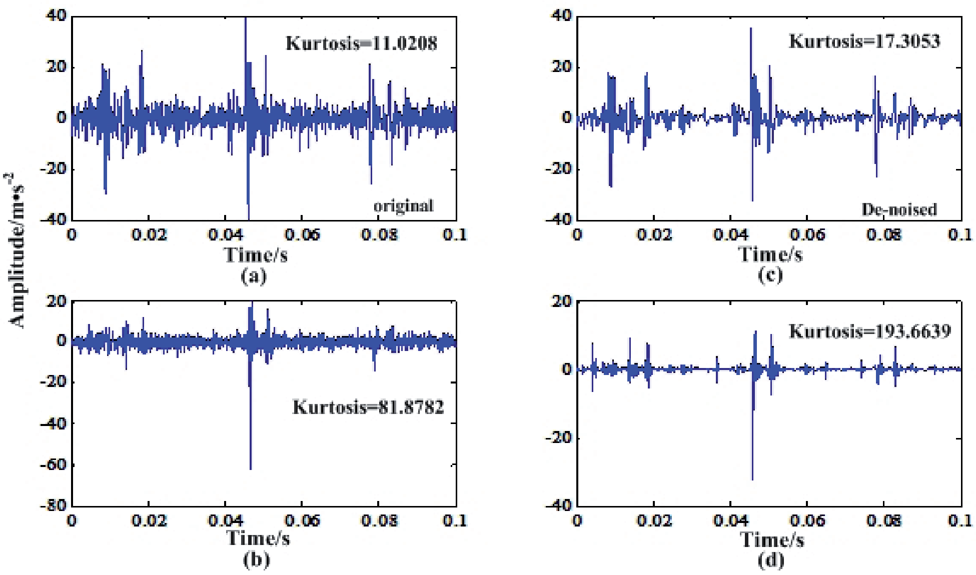


Fig. 6. (a) Original signal; (b) The result of applying AR and MED to original signal; (c) De-noised signal; (d) The result of applying AR and MED to de-noised signal

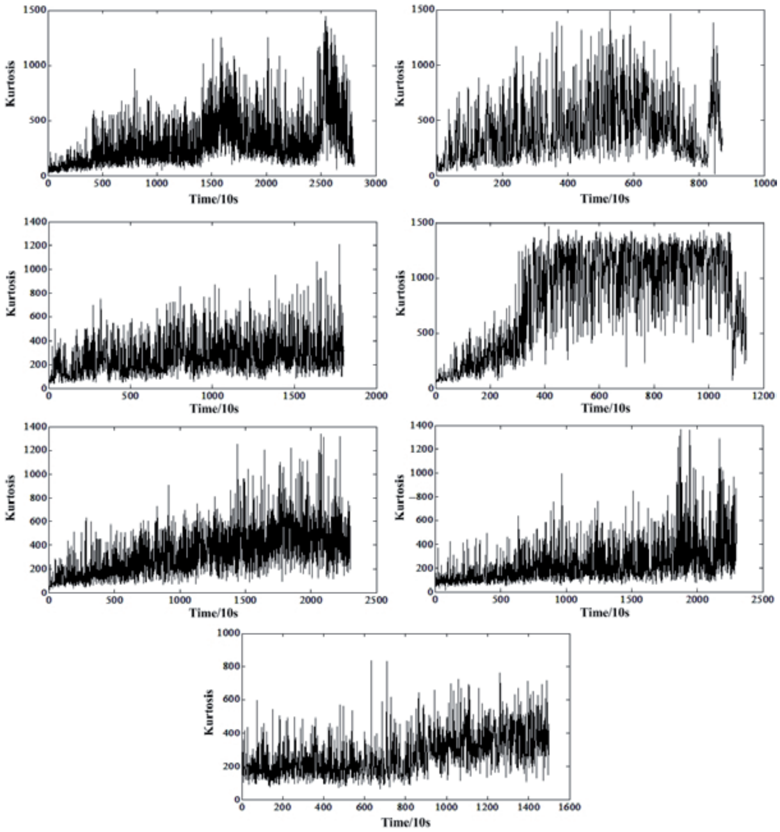


Fig. 7. Features based on AR and MED of 7 bearings

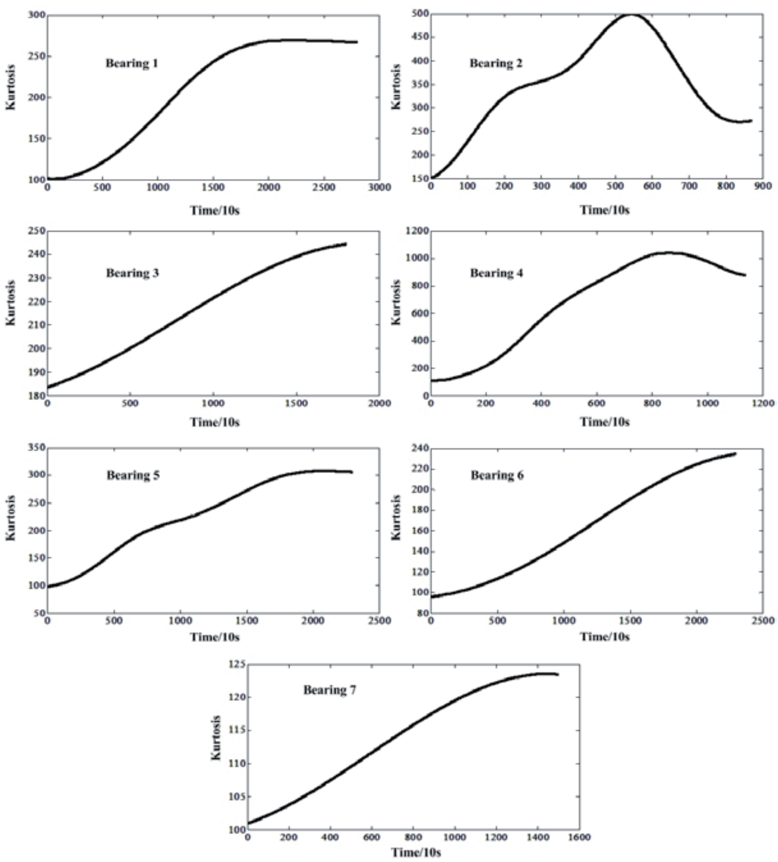


Fig. 8. De-noised features of 7 bearings based on EMD

Figure 8. It can be seen that the extracted signal features clearly show the degradation trend of bearings and enhance considerably the quality of bearing feature extraction. By comparing the de-noised features with RMS and kurtosis values of the bearings (as shown in Figure 4 and Figure 5), inflexion points in the curves of de-noised kurtosis come earlier than those of RMS and kurtosis. It may be inferred that the kurtosis processed using the proposed method can be taken as a leading indicator to evaluate the health condition of bearing, and make a conservative estimation of bearing RUL.

4. Conclusion

In this paper, a hybrid methodology of degradation feature extraction was presented where AR, MED and EMD are employed to enhance the quality of bearing feature extraction and enable bearing prognostics. A case study based on PHM 2012 challenge data was presented to show the performance of the proposed method. In the case study, the features extracted by traditional method fluctuate largely, which can be the barrier for prognostics. It is seen from the case study that the hybrid methodology is effective in degradation feature extraction from bearing vibration signals. Especially, it is applicable to the case that vibration signals fluctuate largely and mask the trend of bearing degradation.

5. References

1. Barszcz T, Sawalhi N. Wind turbines' rolling element bearings fault detection enhancement using minimum entropy deconvolution. Diagnostyka-Diagnostics and structural health monitoring 2011; 59(3): 53–59.
2. Donoho DL, Johnstone I M. Adapting to unknown smoothness via wavelet shrinkage. Journal of American Statistical Association 1995; 90: 1200–1224.
3. Endo H, Randall R B. Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter. Mechanical Systems and Signal Processing 2007; 21(2): 906–919.
4. Grossmann A. Wavelet transform and edge detection. Stochastic Processes in Physics and Engineering 1988; 149–157.
5. Huang NE, Shen Z, Long SR, Wu MLC, Shih HH, Zheng QN, Yen NC, Tung CC, Liu HH. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London Series A – Mathematical Physical and Engineering Sciences 1998; 454(5): 903–995.
6. PHM 2012 bearing challenge data. http://www.femto-st.fr/ieee-PHM2012-data-challenge.
7. Sawalhi N, Randall RB, Endo H. The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. Mechanical Systems and Signal Processing 2007; 21(6): 2616–2633.
8. Wang WY, Wong AK. Autoregressive model-based gear fault diagnosis. Journal of Vibration and Acoustics 2002; 124(2): 172–179.
9. Wang WY, Wong AK. Some new signal processing approaches for gear fault diagnosis. Fifth International Symposium on Signal Processing and its Applications 1999; 587–590.
10. Wiggins RA. Minimum Entropy Deconvolution. Geoexploration 1978; 16: 21–35.

Prof. Hongqiang GU, Ph.D.
Prof. Jianmin ZHAO, Ph.D.
Xinghui ZHANG, Ph.D.
Mechanical Engineering College
Shijiazhuang, Hebei, 050003, P.R. China
e-mail: Jm_zhao@hotmail.com

Article citation info:
PASKA J. Chosen aspects of electric power system reliability optimization. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 202–208.

Józef PASKA

CHOSEN ASPECTS OF ELECTRIC POWER SYSTEM RELIABILITY OPTIMIZATION

WYBRANE ASPEKTY OPTYMALIZACJI NIEZAWODNOŚCI SYSTEMU ELEKTROENERGETYCZNEGO*

Reliability is one of the most important criteria, which must be taken into consideration during planning and operation phases of an electric power system, especially in present situation of the power sector. This paper considers the optimization of electric power system reliability. The formalization of description of electric power system reliability level optimization is done as well as its practical solving components are given: diagram of value based reliability approach and estimation of customer damage costs resulting from insufficient reliability level.

Keywords: *power system reliability, reliability assessment, reliability optimization, outage costs assessment.*

Niezawodność jest jednym z najważniejszych kryteriów, które należy uwzględniać, zarówno podczas planowania rozwoju, jak też eksploatacji systemu elektroenergetycznego, szczególnie w obecnej sytuacji elektroenergetyki. Artykuł dotyczy optymalizacji niezawodności systemu elektroenergetycznego. Przedstawiono formalny opis matematyczny zagadnienia optymalizacji poziomu niezawodności systemu elektroenergetycznego oraz pewne elementy jego rozwiązania: schemat podejścia wartościowania niezawodności oraz szacowanie kosztów strat odbiorców z tytułu niedostatecznego poziomu niezawodności.

Słowa kluczowe: *niezawodność systemu elektroenergetycznego, ocena niezawodności, optymalizacja niezawodności, ocena kosztów przerw w zasilaniu.*

1. Introduction

Reliability of the electric power system (EPS) is defined by its ability to secure the supply of electricity of acceptable quality to the customers [2, 12].

Reliability is one of the most important criteria, which must be taken into consideration during planning and operation phases of a power system [3, 12, 14].

After 1990 deep structural changes occurred and still occur in the electric power systems: takes place disintegration, deregulation and advancing market orientation. This is a worldwide trend. Departure from the vertically integrated structures, deregulation and market solutions in electric industry create new conditions, in which the responsibility for the satisfaction of power demands of individual customers is not and cannot be attributed to the particular electric power company. The objective of the electric power system, which is the assurance of electricity supply of the required quality to the customers at the possibly lowest cost and acceptable reliability of delivery is now the task decomposed into many components, and into many subjects [3].

2. General formulation of the problem of reliability level optimization

Premises of the rational creation of the reliability level should be looked for on the background of economy. Let us use the following terminological convention. Let us divide a set of factors composing the usefulness of the system into two disjoin subsets: the first comprises attributes conditioning the scale, in which the goals of the system may be realized – combination of these attributes values we will call the **productivity of the system**, the second comprises attributes determining reliability level – combination of their values we will call the **system reliability**.

Both productivity and reliability of the system depend on the size, the way and the range of the use of various resources, and funds in the processes of system design, construction and exploitation. The models of resources and funds transformations into a system of specified productivity and reliability form so-called production and reliability functions, and at the same time that transformation comprises processes of laying out and spending funds, material means of various kinds and properties, a human work of various range and qualification level.

A specified system reliability level R^* may be accomplished at many alternative combinations of spending (utilization) of resources. For instance, a specified reliability of a power plant may be achieved by higher investment expenditure (use of better technologies and more expensive materials etc.) or by the higher exploitation costs (skilled and well paid staff, intensive plan of maintenance prevention). Thus a curve of equal (the same) reliability represents all the quantitative combinations of n factors conditioning reliability, resulting with the same effect in the form of reliability level. Neither of those combinations is better than the others, if we mean the final result, and the choice of optimal combination of resources is conditioned by two factors:

- Relative effectiveness of individual resources and/or methods of their utilization (in the sense of influence on reliability),
- Relative value or cost of individual resources and/or the ways of their utilization.

To describe the problem of optimal reliability there must be also introduced the concept of marginal reliability R' in relation to resources $\mathbf{X}$ (where $\mathbf{X} = \{X_1, X_2, ..., X_n\}$ is a vector of resources - funds) and considered valuable aspect of resources transformation into reliability.

Marginal reliability R' in relation to resources $\mathbf{X}$ describes the changes in system reliability R, when during its design, construction and exploitation units of individual resources are added or subtracted. Thus, in a specified point:

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

$$R_j^{'} = \frac{\delta R}{\delta X_j} \quad (1)$$

where: $R_j^{'}$ – marginal reliability in relation to j-th resource,
R – reliability expressed by its measure (reliability index) of physical character (for instance undelivered energy, frequency of failures in supply),
X_j – j-th resource;

or, when the partial derivatives do not exist:

$$R_j^{'} = \frac{\Delta R}{\Delta X_j} \quad (2)$$

In majority of cases marginal reliability $R_j^{'}$ decreases when X_j increases, which means that measured in categories of reliability marginal effect (product) of any resource decreases with the growth of quantity of utilized (spent) resource and remaining without changes quantities of the rest of resources. For example, in a power station the growth of funds on staff qualification level improvement, without modernization of equipment and/or raise of funds on planned maintenance, received increments of reliability will be smaller and smaller.

Valuable aspect of resources transformation is included with the help of suitable valuable model. Thus we have two models:

1. **Physical model**, in principle considered until now, in which we have:
   - Quantity of spent and/or utilized resources, $\mathbf{x} = \{x_1, x_2, ..., x_n\}$, where $\mathbf{x}$ is realization of resources vector $\mathbf{X}$;
   - Reliability represented by its measure (reliability index) of physical character, R;
   - Reliability function, $R = r(\mathbf{X})$;
2. **Valuable model**, in which we have:
   - Value of resources spent and/or utilized to ensure reliability, $V(\mathbf{X})$;
   - Reliability value, $V(R)$ or economic effect of its unsatisfying level (economic and/or social losses), $SL(R)$;
   - Objective function, determining value (profit) or cost of transformation of resources $\mathbf{X}$ into reliability R,

$$P = V(R^*) - V(\mathbf{X}) \to \max \quad (3)$$

$$C = V(\mathbf{X}) + SL(R) \to \min \quad (4)$$

with the constraints of the type:

$$R' \le r(\mathbf{X}). \quad (5)$$

The task described by the objective function (3) is a task of optimal choice, in the aspect of specified reliability level R^*, variant of spending and/or utilizing resources $\mathbf{x}$, and task (4) looks for optimal system reliability R.

Optimization tasks formulated by relations (3) – (5) have solutions, when values of reliability and resources can be measured by the same units, for instance by monetary units. If it is not like that, one of the following problems may be solved:

1. Minimization of values of spent and/or utilized reserves to achieve desired reliability level.

$$C = V(\mathbf{X}) \to \min , \quad (6)$$

with the constraints:

$$R^* = r(X_1, X_2, ..., X_n). \quad (7)$$

2. Maximization of reliability level with given or limited resources.

$$R = r(\mathbf{X}) \to \max, \quad (8)$$

with the constraints:

$$C^* = V(X_1, X_2, ..., X_n). \quad (9)$$

Using the method of Lagrange's multipliers the following conditions of extreme existence are received:

- for problem 1

$$\frac{R_i^{'}}{C_i^{'}} = \text{const.}, \quad \text{for } i=1, 2, ..., n ; \quad (10)$$

- for problem 2

$$\frac{C_i^{'}}{R_i^{'}} = \text{const.}, \quad \text{for } i=1, 2, ..., n ; \quad (11)$$

where: $C_i^{'}$ - marginal cost of i-th resource, understood as valuable measure of the increment of spending and/or utilizing of the i-th resource to achieve growth of reliability level by one unit.

Thus optimal reliability level is determined by a point, in which the ratios of marginal reliabilities to marginal costs are equal. For simplicity one representative measure for reliability R has been adopted here, but there is no obstacle to determine reliability level by the values of several indices.

Requirements for continuity, convexity, rationality and comparability of reliability function and spending and/or utilizing of resources function are very difficult to fulfill in practice. But successfully one may consider the problem of optimal, from the reliability point of view, expending of limited resources to satisfy certain needs. In that case we can distinguish three variants of the problem:

1. At defined quantity of resources and at given technical constraints we should maximize system reliability.
2. At defined, required system reliability level and at given technical constraints we should minimize expenditure and/or utilization of reserves necessary to gain and keep reliability.
3. It is necessary to achieve such combination of reliability and resources usage to reach it and keep it (reliability), which maximize the degree of system realization goals.

Real problems of system reliability optimization belong to the uncertain class of problems (rarely probabilistic) multi-dimensional and compound, dynamic and multi-criteria.

3. Optimal reliability of electric power system

The task of electrical power system is to ensure supply of electrical energy to the customers with required quality and at the possible lower cost and acceptable reliability of delivery. In this case, also, the cost of ensuring a certain level of reliability supply should be concerned to the value of reliability for a customer.

For a power system the relations (3) and (4) may be written as:

$$P(A, R) = VS(A, R) - CoS(A, R) \to \max \quad (12)$$

$$C(A, R) = CoS(A, R) + COC(A, R) \to \min \quad (13)$$

where: $P(A, R)$ – social value (profit) of covering request (demand) on electrical energy A with reliability R,
$VS(A, R)$ – the value of energy sale at the quantity of A with the reliability of R (it is the inclination of a customer for paying to use A energy at its reliability delivery R),

$CoS(A, R)$ – costs of covering the demand A with reliability R,
$COC(A, R)$ – losses costs coming from insufficient reliability R,
$C(A, R)$ – total (social) cost of covering the demand A with reliability of R.

In relations (12) and (13) the demand A is a function of reliability R, and economic quantities are annual values or the sums of discounted values for the whole long-term period of analysis.

From the necessary condition for the existing of extreme it comes that:

$$\frac{dP}{dR} = 0 \rightarrow \frac{dVS}{dR} = \frac{dCoS}{dR} \qquad \frac{dC}{dR} = 0 \rightarrow \frac{dCoS}{dR} = -\frac{dCOC}{dR} \tag{14}$$

which means that at optimal reliability there are equalities of two corresponding marginal economic values.

In practice the relation between the demand for electrical energy and reliability of its delivery R is usually neglected (non-flexibility of demand against reliability is assumed). Assuming that the reliability level R is represented by an index ensuring energy delivery EIR (Energy Index of Reliability), the illustration of relation (13) is shown in Fig. 1.

Higher reliability level R needs rising costs of ensuring that reliability – costs of "deliverer-supplier", but in result it gives decreasing cost of widely understood losses at "customer", coming from insufficient reliability. Comparison of these two economic categories leads to the definition of "optimal" reliability level or optimal value of, representative for reliability (in the given analysis), quantity characterizing power system as a whole or its subsystem (for example reserve or capacity margin).

Optimal reliability level R_{opt} means minimal total cost C:

$$R_{opt} = R\colon \ \frac{dC}{dR} = 0 \rightarrow \frac{dCoS}{dR} = -\frac{dCOC}{dR} \tag{15}$$

and it does not cover with the reliability level, at which the equalization of the cost of reliability insurance takes place, $R_=$

$$R_= = R\colon \ CoS = COC \tag{16}$$

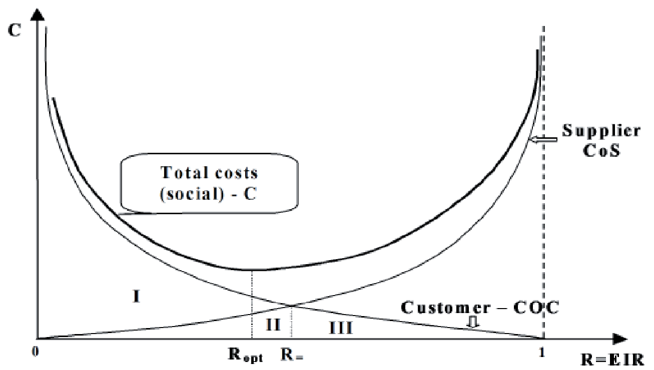


Fig. 1. Complete (social) reliability costs: C – cost; R – reliability level, represented by index EIR; R_{opt} – optimal reliability level; $R_=$ – reliability level, at which there exists equalization between the costs of ensuring reliability with the cost of losses, coming from its (reliability) insufficient level; I – effective area for actions to improve reliability; II – intermediate area; III – non-effective area for actions to improve reliability

In Fig. 1 we can distinguish three areas: I – area of action effectiveness to improve reliability, in which these actions give a result decreasing total cost and the rate of losses cost drop is higher than the velocity of growing costs of ensuring reliability, II – intermediate area, in which the total cost grows slightly, III – non-effective area for actions to improve reliability with the higher and higher rise of the total cost.

The goals of actions in the area of electric power system reliability are the following:

- Keeping the existing level of system reliability.
- Identification of investment projects which give the most important share into ensuring or improving system reliability.
- Defining and denomination of quantitative measures (indices) of reliability for the purposes of system development planning.
- Ensuring that the system parameters in the future will fulfill requirements of its reliability.
- Valuation of reliability in categories of costs of losses caused by the breaks and limitations of electrical energy supply.

Assessment of economic losses caused by unreliability of electric power system is, in particular, necessary for analyzing alternative plans of grid systems development. Advisability of undertaking investments rising system reliability may be assessed on the background of relation of costs and forecasted benefits. The tool in such understood system development planning is analysis cost – benefit, known as a reliability valuing (VBRA - Value-Based Reliability Approach) [19].

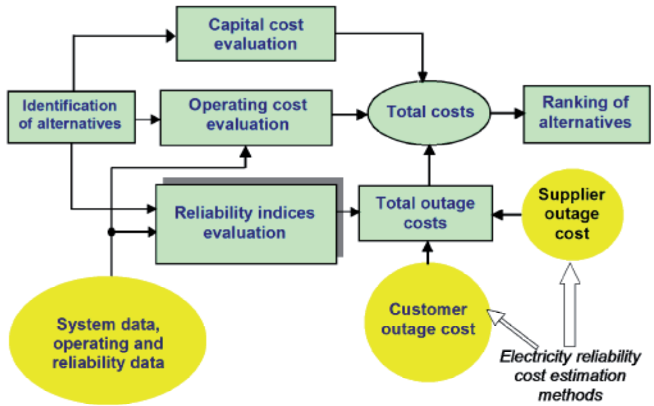


Fig. 2. General idea of reliability valuation

Basic components of reliability valuing are (Fig. 2):

- Identification of alternative projects.
- Assessment of capital and operational costs (connected with actions keeping or rising system reliability).
- Calculation of reliability indices for planned system structures.
- Assessment of costs of losses caused by interruptions and limitations of electricity supply.
- Ranking of alternative projects taking into account their total cost of solution.

Total costs are used to classify the variants of development or exploitation of a system. The total (discounted) cost is given by a formula:

$$C = CoS + COC = C_I + C_o + COC \tag{17}$$

where: C – total cost of a variant,
C_I – investment expenditure of a variant,
C_o – operational costs (exploitation),
COC – undelivered energy cost (customer outage cost).

Thus we are looking for a variant, which has got minimal costs including costs of losses at customers, caused by breaks and limitations, and necessary investment expenditures and exploitation costs in the whole perennial period of exploitation.

4. Costs of unreliability and their estimation

Important component that valuate variants of extending, modernization and exploitation of a power system are losses coming from breaks and limitations of electrical energy supply to customers, which were defined in equation (17) as the cost of undelivered energy [4, 5, 10]. These losses are difficult to estimate because there is no simple relationship between undelivered energy and economic losses (harms), which a customer is going to take. It depends on many factors; to the most important we can include varying intensity of a customer action. For example it may be in an industry – phase of technological process, kind of a shift, season of a year; in trade – intensity of buying and selling; at home – connection with other outside factors, time of break and so on. Thus we can say that there is not always strong correlation between undelivered energy and customer's economic losses. The same value of undelivered energy in different periods of firm's work may cause various economic losses. Those economic losses should be represented by a value of undelivered energy estimated by, e.g. a customer in public opinion poll, done in a wide scale (Fig. 2). A customer answers a range of questions and particularly the following question: how much he would pay to avoid a break in supply at given conditions? Value of undelivered energy usually settled in this way composes its marginal value for a break of defined duration, in given work conditions and situation of an enterprise – in other conditions losses at the same undelivered energy may be completely different [1, 6, 7, 8].

Despite of those reservations there is a need for such a rough measure as Interrupted Energy Assessment Rate (IEAR) – in Polish bibliography called economic equivalent of undelivered electrical energy, denoted by k_a [17], which multiplied by undelivered energy gives the assessment of economic losses.

Sufficiency of employing of this index comes from the fact that each improvement of power system reliability goes by leaps. For example change from unilateral feeding of a customer to double-sided one causes improvement of reliability in order of several dozen times. It means meaningful softening of requirements as to the precision of estimation of economic losses caused by breaks in supply – in many cases only approximate values are enough.

Investigations and analysis carried for groups of customers give information on costs of "interruption/outage" and not on costs of "kW interrupted or curtailed power", or on costs of "kWh undelivered energy". Then they are transformed into the form "cost/kW" or "cost/kWh" and given for distinguished groups of customers and characteristic values of interruption duration. For instance, in investigations conducted by University of Saskatchewan for Canada in eighties, there were distinguished 7 groups of customers: large customers, industry, commerce and services, agriculture, household, government institutions and public utilities, offices and buildings; and 5 characteristic values of interruption duration: 1 min., 20 min., 1 h, 4 h, 8 h. Received values determine so-called SCDF and may be used for analysis on the third hierarchical level of a power system (generation, transmission and distribution together) – HL III.

Newer studies were made in Great Britain on the area of three distribution companies (Manweb, MEB, Norweb) in the period from October 1992 to March 1993. In their effect SCDF was determined for four distinguished groups of customers: household (residential), commerce and services (commercial), industry, large users (above 8 MW) and seven characteristic values of interruption duration: momentary break, 1 min., 20 min., 1 h, 4 h, 8 h, 24 h. They are presented in Table 1.

The relation between $SCDF_A$ and $SCDF_P$ from Table 1 is as follows:

$$SCDF_A = \frac{SCDF_P}{8760m}, \tag{18}$$

where: m is average annual degree of sector load (load factor).

Extensive research, funded by the Department of Energy have been made in the United States, in the years 1989-2005. The synthetic results are summarized in Table 2. They provide an estimate obtained from the analysis of the results of 28 surveys carried out by 10 major U.S. energy companies, which included 11 970 firms and 7 963 households. The values in Table 2 are averaged, independent of the time of interruption (time of year, working day or holiday, the time of day).

To make analysis on the second hierarchical level of a power system – HL II (generation and transmission together) we must have CCDF, which determines costs of losses, as a result of interruptions and curtailments, of customers from certain area [$/kWh, zł/kWh] as a function of failure duration. To construct such function we must have SCDF functions for specified groups of customers and the share of these groups in energy demand. We can also make aggregation and determine costs falling on a failure and on undelivered energy unit.

$$CCDF_A = \sum_{s\in ns} \frac{SCDF_{P,s}}{m_s 8760}\left(\frac{A_s}{\sum\limits_{s\in ns} A_s}\right) = \sum_{s\in ns} SCDF_{A,s}\left(\frac{A_s}{\sum\limits_{s\in ns} A_s}\right) \tag{19}$$

where: s – sector (group) of customers supplied from considered subsystem (node, bus),
ns – number of customer sectors in considered area,
m_s – average annual degree of sector load,
A_s – annual energy consumption by sector s.

Table 1. SCDF in Great Britain (UK, 1992) [9]

Duration	Household		Commerce & Services		Industry		Large customers	
	$SCDF_A$	$SCDF_P$	$SCDF_A$	$SCDF_P$	$SCDF_A$	$SCDF_P$	$SCDF_A$	$SCDF_P$
Momentary	-	-	0.46	0.99	3.02	6.15	1.07	6.74
1 min.	-	-	0.48	1.02	3.13	6.47	1.07	6.74
20 min.	0.06	0.15	1.64	3.89	6.32	14.27	1.09	6.86
1 h	0.21	0.54	4.91	10.65	11.94	25.26	1.36	7.18
4 h	1.44	3.72	18.13	39.04	32.59	72.22	1.52	8.86
8 h	-	-	37.06	78.65	53.36	120.11	1.71	9.71
24 h	-	-	47.58	99.98	67.10	150.38	2.39	13.35

$SCDF_A$, GBP/kWh – costs of losses per kWh of energy used during a year by an average customer of a sector;
$SCDF_P$, GBP/kW - costs of losses per kW of peak demand by an average customer of a sector.

Table 2. Estimating the average cost of losses caused by power outages in the U.S. (in USD_{2008}) [7]

Outage costs	Duration of interruption				
	< 5 min.	30 min.	1 h	4 h	8 h
Medium and large commercial and industrial customers*					
Cost per event, USD	6 558	9 217	12 487	42 506	69 284
Cost per kW average power demand, USD/kW	8	11.3	15.3	52.1	85
Cost per kWh of undelivered energy, USD/kWh	96.5	22,6	15.3	13	10.6
Cost per kWh of energy consumed in a year, 10^{-3} USD/kWh	91.8	1.29	1.75	5.95	9.7
Small commercial and industrial customers**					
Cost per event, USD	293	435	619	2 623	5 195
Cost per kW average power demand, USD/kW	133.7	198.1	282	1 195.8	2 368.6
Cost per kWh of undelivered energy, USD/kWh	1 604.1	396.3	282	298.9	296.1
Cost per kWh of energy consumed in a year, USD/kWh	0.00153	0.00226	0.00322	0.137	0.27
Households					
Cost per event, USD	2.1	2.7	3.3	7.4	10.6
Cost per kW average power demand, USD/kW	1.4	1.8	2.2	4.9	6.9
Cost per kWh of undelivered energy, USD/kWh	16.8	3.5	2.2	1.2	0.9
Cost per kWh of energy consumed in a year, 10^{-4} USD/kWh	1.6	2.01	2.46	5.58	7.92

** - with the annual consumption of more than 50 MWh, ** - with an annual energy consumption not exceeding 50 MWh*

In case when there are problems to determine the degree of load value for each sector we can use the simplified relation:

$$CCDF_A = \sum_{s\in ns} SCDF_{P,s}\left(\frac{A_s}{\sum_{s\in ns} A_s}\right)\frac{1}{m8760}, \tag{20}$$

where m is the degree of load (load factor) of the considered area.

A global index, known as IEAR, and in Poland known as economic equivalent of undelivered electrical energy is used in power system analysis at levels HL I and HL II. It has dimension USD/kWh, zł/kWh and multiplied by an expected value of undelivered energy (*LOEE, EENS*, *EUE*) gives assessment of social costs of losses caused by inadequate reliability level.

In similar character, but with other name, parameter IEAR is used in models for planning development and for assessment of generation subsystem (WASP III - ELECTRIC, ICARUS, IPM, PLEXOS) and it is met in expressions on marginal costs.

On the other hand the parameter appearing until the year 2000 in expression on the price of energy purchase from generators in the electricity pool in England and Wales – VOLL[1] and the parameter appearing in expression on unitary offer price paid to generators in the project of Polish offer system of electrical energy market [20] - KNZ[2] are the measure giving, assessed by the customers, value of electrical energy in the situation of its lack. It is therefore marginal price, which a customer would be eager to pay in extreme conditions.

To find IEAR we can use functions CCDF or SCDF. The process may be brought to two steps:

Step 1: Determination of the IEAR as a function of interruption duration - IEAR(t_p).

$$IEAR(t_p) = \frac{CCDF(t_p)}{t_p m} \tag{21}$$

or

$$IEAR_s(t_p) = \frac{SCDF(t_p)}{t_p m_s} \tag{22}$$

and then

$$IEAR(t_p) = \sum_{s\in ns} IEAR_s(t_p) k_s \tag{23}$$

where: k_s – weighted coefficient, for example: relative annual energy consumption (the best), relative number of customers, relative peak load.

Step 2: Determination of expected value of the IEAR.

$$IEAR = \sum_{t_p=0}^{T_{gr}} IEAR(t_p)\mathrm{p}(t_p) \tag{24}$$

where: $\mathrm{p}(t_p)$ – probability of existence of interruptions lasting t_p,
T_{gr} – maximum time in which supply should be restored, e.g. it is required in Great Britain that the supply should be restored during 24 hours – in Poland from the year 2007 it is also 24 h, according to "system regulation" [15].

It is possible to use approximate value, which is the ratio of gross domestic product (GDP) to the total electrical energy consumption (EE) – in "Polish power industry statistics" [18] the inverse of that relation is given (EE/GDP). Values of economic equivalent – index of undelivered energy value (k_a – *IEAR*), determined in such a way, are listed in Table 3 (in zł/MWh and prices from the year 2005).

According to "tariff's regulation" [16] a customer in Poland has right to compensation for each unit of undelivered electrical energy, equal to tenfold (customer connected to the network with voltage not higher than 1 kV) or fivefold (rest of customers) price of energy for

1 Value of Lost Load. Its initial value of 2 GBP/kWh adopted in 1989 by Electricity Supply Industry has been increased with *RPI* (Retail Price Index). Despite no objective reasoning for adopting this figure *VOLL* has had a very important role in the setting of England and Wales pool payments (Capacity Element of Pool Purchase Price).

2 Cost of not covered demand (it should rather be: value of not covered demand). *KNZ* was established in 1997 as 1.45 zł/kWh.

Table 3. Approximate values of undelivered energy economic equivalent (GDP/EE) in Poland

Year	1995	2000	2004	2005	2006	2007	2008	2009	2010
k_a, zł/MWh	4138	5125	5767	5922	6925	7220	7570	7968	7918

the period, in which interruption took place. So it was assumed that there is proportional relationship between the values of losses caused by energy supply interruption and the quantity of undelivered energy. The cost of these losses, falling on the unit of undelivered energy, was assessed to be ten or five times higher that the energy price.

Losses occurring at the customer, resulting from interruptions and limitations of the supply of electricity are difficult to estimate because of the lack of a stable model of relationship between the undelivered energy and the losses suffered by the customer. One of the applied methods is the survey method, which allows the estimation of losses caused by not supplying electricity to customers and to obtain information about the worth of energy supply reliability for the customer.

A commonly accepted method of evaluating customer interruption costs is to directly survey electric utility customers. The basis of this approach is that customers are in best position to understand power interruption and limitations in power quality consequences in relation to their particular needs.

One of the goals of "TRELSS for the PPGC" project [13] was to create a base for the use of reliability evaluation approach at the Polish Power Grid Company. Special questionnaire, based on the experience of EPRI [5, 11], was prepared to realize that aim.

Customers were divided into two groups: residential and industrial/commercial. For each scenario of an interruption a customer is asked in the questionnaire to answer a question what would happen in his house (for residential sector) or in his firm (for industrial or commercial sector) as a result of the interruption and what his reaction is on that event. The customer is asked to evaluate that presumable event in zlotys (Polish currency). The value in zlotys may be clearly defined in one of the following three cases: direct costs, tendency to payment or desire for taking recompense. In the first case the customer is asked to determine direct economic costs of his activity (business), which was subjected to losses because of the interruption. Those costs may comprise: lost production costs and inactivity costs. Description of consequences of interruptions was enclosed to help the customer to think over all the cost components and choose the proper ones. In the second case the customer is asked to choose maximal value of payment from among given in the questionnaire that he would accept to avoid such event. In the third case the customer is asked to choose minimal value from among given in the questionnaire that he would accept as recompense for the cost of the event. Typical inclination to payment is defined as a fraction of the determined direct cost. The customer must decide himself, which of the given values he would use.

Customers are also asked in the questionnaire to give some information about themselves, their firms or families and, what is the most important, about the latest experiences with the interruptions and their service, i.e. liquidation of the interruptions. This information will be used in econometric analysis to identify factors, which have the greatest influence on costs of interruptions.

Residential customers

Residential (household) customers fullest recognize the importance of electricity in his apartment/house when an unexpected supply interruption or distortion of quality, such as flickering lights, voltage dips, etc. During the interruption the customers do not have access to electrical appliances. This forces them to change daily habits, schedule of activity and usually puts in a forced, painful situation. It may specify:

- coercion idle for lack of ability to perform normal household duties and any additional paid work, which (usually, or sometimes) we make at home;
- sometimes difficulties for children who cannot learn or make up your homework;
- impediment to rest;
- deterioration in comfort staying at home, for example, the simultaneous cessation of home heating in the winter, because it is dependent on the energy supply, the lack of ventilation or air conditioning in the summer, the lack of water, etc.;
- threat to health and life, when one member of the household is ill or requires constant care or medical apparatus's work;
- need to tune some devices (clocks, computers, protections, etc.) after restoring the supply of electricity;
- food spoilage, damage to other equipment, etc.;
- other problems.

Willing to pay / willingness to accept payment is a useful measure for assessing the cost of interruptions because the residential (household) customers have difficulty in estimating the direct costs of outages, especially unexpected. The questionnaire for the residential customers has about 50 questions divided into 5 sections.

Industrial/commercial customers

For these users the test of the cost of interruptions is made for power interruptions hypothetical situations. Cost of break can be determined by comparing two scenarios:

1. scenario **with interruption**, includes duration of interruption, time for interruption consequences' clearing until the conditions are reached close to existing before the break,
2. scenario **without interruption**, so-called "motion" normal, at the same time as in scenario 1 (with interruption).

The questionnaire for industrial/commercial customers contains about 80 questions divided into 6 sections.

Poll investigations of losses' costs, resulting from interruptions in electrical energy supply, carried until now may be divided into two phases:

- Test phase, checking the questionnaire itself and its acceptance by customers, which had been done during realization of the project "TRELSS for the PPGC" [13].
- Active phase, in which after some slight corrections in the questionnaire, investigations were carried on greater, but still not big enough, sample of customers, in two regions of Poland: in central part of Poland and on the south of the country.

Unfortunately, due to lack of interest from power utility companies and the regulatory authority, more extensive research on the cost of losses due to interruptions in the electricity supply, in the current realities of the functioning of the Polish power sector and the economy, have not yet been carried out.

5. Conclusions

Reliability of an electric power system decides about the quality of supply and consumers' trust, that they will get energy adequate to their requirements. Even, or rather particularly, in the present situation of the liberalization of electricity markets and unbundling of generation, transmission and distribution, questions on the present and future reliability level arise, and the interest in detailed investigation of electric power system reliability issues, especially taking into account possibly the whole power system, increases.

Real problems of power system reliability optimization belong to the class of uncertain problems (rarely probabilistic), multi-dimensional and compound, dynamic and multi-criteria. Their practical solution requires acceptance of far reaching simplifying assumptions.

The players in today's electricity market view economic processes in a short-term scale. However, improving supply reliability and power system reliability requires the management of longer-term financial and/or physical risks, which requires longer-term incentives. They could be justified only basing on the results of cost-benefit reliability analysis – reliability optimization.

References

1. Assessment of the Value of Customer Reliability. CRA International, 12 August 2008.
2. Billinton R, Allan RN. Reliability Evaluation of Power Systems, second ed. New York: Plenum Press, 1996.
3. Billinton R, Salvaderi L, McCalley JD, Chao H, Seitz Th, Allan RN, Odom J, Fallon C. Reliability issues in today's electric power utility environment. IEEE Transactions on Power Systems 1997; 12: 1708–1714.
4. Burns S, Gross G. Value of service reliability. IEEE Transactions on Power Systems 1990; 5: 825–834.
5. Cost-Benefit Analysis of Power System Reliability: Determination of Interruption Costs. TR-2878. Palo Alto: EPRI, 1990.
6. Cost of Power Interruptions to Electricity Consumers in the United States (U.S.). Ernest Orlando Lawrence Berkeley National Laboratory, LBNL-58164, February 2006.
7. Estimated Value of Service Reliability for Electric Utility Customers in the United States. Ernest Orlando Lawrence Berkeley National Laboratory, LBNL-2132E, June 2009.
8. Investigation into the value of lost load in New Zealand – Summary of findings. Electricity Authority, 13 January 2012.
9. Kariuki KK, Allan RN. Evaluation of reliability worth and value of lost load. IEE Proc. – Generation, Transmission, Distribution 1996; 143: 171–180.
10. Methods to Consider Interruption Costs in Power System Analysis. Report No. 191,Task Force 38.06.01, CIGRE, August 2001.
11. Outage Cost Estimation Guidebook. TR-106082. Palo Alto: EPRI, 1995.
12. Paska J. Niezawodność systemów elektroenergetycznych. Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej, 2005.
13. Paska J, Bargiel J, Bartczak J, Goc W, Kłos A, Momot A, Nowakowska E, Sowa P, Teichman B. Application of TRELSS and Implementation of Value-Based Transmission Reliability Approach at Polish Power Grid Company. TR-114816. Palo Alto: EPRI, March 2000.
14. Paska J, Bargiel J, Oleksy A. Application of Value-Based Reliability Approach in Power Transmission System Planning. 7th International Conference on Probabilistic Methods Applied to Power Systems – PMAPS 2002; Naples - Italy, September 22–26, 2002.
15. Rozporządzenie Ministra Gospodarki z dnia 4 maja 2007 w sprawie szczegółowych warunków funkcjonowania systemu elektroenergetycznego. Dz. U. 2007 r. Nr 93, poz. 957; 2008 r. Nr 30, poz. 178; 2008 r. Nr 162, poz. 1005.
16. Rozporządzenie Ministra Gospodarki z dnia 18 sierpnia 2011 w sprawie szczegółowych zasad kształtowania i kalkulacji taryf oraz rozliczeń w obrocie energią elektryczną. Dz. U. Nr 189, poz. 1126.
17. Sozański J. Niezawodność i jakość pracy systemu elektroenergetycznego. Warszawa: WNT, 1990.
18. Statystyka elektroenergetyki polskiej 2010. Warszawa: Agencja Rynku Energii, 2011.
19. Value-Based Transmission Resource Analysis. Research Project 2878-02. Palo Alto: EPRI, April 1994.
20. Zasady działania systemowego rynku ofertowego – pool'u. Warszawa: Energoprojekt-Consulting S.A., 1997.

Prof. Józef PASKA, Ph.D., D.Sc. (Eng.)
Warsaw University of Technology, Institute of Electrical Power Engineering,
Group of Electric Power Plants and Economics of Electrical Power Engineering
ul. Koszykowa 75
00-662 Warszawa, Poland
e-mail: Jozef.Paska@ien.pw.edu.pl

INFORMATION FOR AUTHORS

Terms and Conditions of Publication:

- The quarterly „Maintenance and Reliability" publishes original papers written 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.

Fees:

- Pursuant to a resolution of the Board of PNTTE, as of 2009 the publication fee for one text is 600 zloty + VAT.
- Coloured graphical elements in the submitted text require agreement from the Editor and are charged extra.

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 graphical files in the *.tif, *.jpg or *.cdr (v. X3) formats.
- A manuscript should include (in accordance with the enclosed correct manuscript format: *.pdf, *.doc):
- 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.

Editor contact info: (Submissions should be sent to the Editor's address)

Editorial Office of „Eksploatacja i Niezawodność - 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.