Science and Technology

VALIŠ D, VINTR Z, MALACH J. **Selected aspects of physical structures vulnerability– state-of-the-art.** Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14 (3): 189-194.

The paper is to deal with the selected aspects of structures vulnerability in terms of the physical protection. It is going to specify possible risks following from a terrorist attack, define a potential target and its characteristics, describe the resistance of an object to an attack, and determine the ways to reduce the probability of reaching a terrorist target, or increase object resilience. The results we are going to introduce reflect current knowledge in the area of physical protection.

VALIŠ D, VINTR Z, MALACH J. **Wybrane zagadnienia dotyczące narażenia obiektów fizycznych na atak – stan wiedzy.** Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14 (3): 189-194.

W artykule omówiono wybrane aspekty narażenia obiektów fizycznych na atak w kontekście ochrony fizycznej. Określono możliwe zagrożenia wynikające z ataku terrorystycznego, zdefiniowano potencjalne cele ataku i podano ich charakterystyki, opisano zjawisko wytrzymałości obiektu na atak, oraz ustalono sposoby zmniejszania prawdopodobieństwa dotarcia do celu ataku terrorystycznego lub zwiększania odporności obiektu. Przedstawione wyniki odzwierciedlają bieżący stan wiedzy w dziedzinie ochrony fizycznej.

PANG Y, HUANG HZ, XIAO NC, LIU Y, LI YF. **Possibilistic Reliability Analysis of Repairable System with Omitted or Delayed Failure Effects**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 195-202.

Within the practical problems in industrial engineering, the failure effect sometimes can be omitted or delayed if it has less effect on the system. In detail, the prominent features of the system can be described as follows: 1) if a repair time is sufficiently short (less than some threshold value) that does not affect the system operation, i.e. the pessimistic effect of system failure could be ignored. The system can be considered as operating during this repair time. It is called the system with repair time omission (failure effect omitted). 2) if a repair time is longer than the given threshold value and the failure effect is finally suffered. Then the system can be considered to remain operating from the initial stage of the repair till the end of the repair threshold. It is called the system with delayed failure effect. Based on the above two characteristics, model for the related repairable system is introduced in this paper. Two scenarios are discussed where the threshold value is regarded as a constant and non-negative random variable, respectively. Reliability indices such as instantaneous possibilistic availability are formulated for the system with failure effect omitted or delayed.

PANG Y, HUANG HZ, XIAO NC, LIU Y, LI YF. **Posybilistyczna analiza niezawodnościowa systemu naprawialnego z pominiętym lub opóźnionym efektem uszkodzeniay**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 195-202.

Przy rozwiązywaniu problemów praktycznych w inżynierii przemysłowej można czasami pominąć bądź opóźnić efekt uszkodzenia jeśli ma on niewielki wpływ na system. Ściślej, wiodące cechy systemu można opisać w następujący sposób: 1) jeżeli czas naprawy jest wystarczająco krótki (krótszy niż pewna wartość progowa), tak iż nie ma on wpływu na działanie systemu, to można pominąć negatywny efekt uszkodzenia systemu. Przy takim czasie naprawy można uznać że system nie przerwał działania. Nazywa się go wtedy systemem z pominięciem czasu naprawy (pominięty efekt uszkodzenia). 2) Jeżeli czas naprawy jest dłuższy niż dana wartość progowa i efekt uszkodzenia staje się w końcu odczuwalny, to uznajemy, że system pozostawał aktywny od początkowego etapu naprawy aż do momentu, w którym został przekroczony próg czasu naprawy. Nazywa się go wtedy systemem z opóźnionym efektem uszkodzenia. W oparciu o powyższe dwie charakterystyki, wprowadzono w prezentowanej pracy model systemu naprawialnego. Omówiono dwa scenariusze, w których, odpowiednio, przyjęto, że wartość progowa jest wartością stałą lub nieujemną zmienną losową. Sformułowano wskaźniki niezawodnościowe, takie jak posybilistyczna gotowość chwilowa, dla systemu z pominiętym lub opóźnionym efektem uszkodzenia.

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.

Methodology for determining the limiting time of the hot water-pipeline exploitation has been described in the present study. Pitting corrosion causing local reduction of the hot water-pipeline wall thickness was assumed as the basis of the limiting time determining. Three limiting states influencing the hot water-pipeline strength were taken into consideration in the executed analysis. It was estimated that the wall thickness distribution is consistent with normal distribution as well as that exceeding of the hot water-pipeline wall thickness limiting values within given probability level is a basis for the exploitation time determining.

PILCH R, SZYBKA J, BRONIEC Z. **Wyznaczanie czasu eksploatacji ciepłociągu na podstawie identyfikacji stanów granicznych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 203-207.

W artykule przedstawiono metodykę wyznaczania granicznego czasu eksploatacji rurociągu. Podstawą wyznaczenia tego czasu jest korozja wżerowa powodująca lokalne zmniejszanie grubości ścianek rurociągu. W analizie wzięto pod uwagę trzy stany graniczne decydujące o wytrzymałości rurociągu. Oszacowano, że rozkład grubości ścianek jest zgodny z rozkładem normalnym a przekroczenie granicznych wartości grubości ścianki rurociągu na zadanym poziomie prawdopodobieństwa jest podstawą do określenia czasu jego eksploatacji.

LI YF, HUANG HZ, LIU Y, XIAO N, LI H. **A new fault tree analysis method: fuzzy dynamic fault tree analysis**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 208-214.

Fault tree analysis (FTA) is a widely used as a reliability assessment tool for large and complex engineering systems. The conventional fault tree analysis method, which contains AND, OR, and Voting gates, etc., can efficiently build an analytical model to represent combinations of component failures that cause the failure of a system. However, due to its limited modelling capability, we may confront difficulties when modelling dynamic systems which involve complicated dynamic characteristics such as sequence dependency and functional dependency. Markov-based dynamic fault tree analysis (DFTA) extends the static FTA by introducing additional gates to model such complicated interactions among events. In many circumstance, it is quite difficult to obtain an accurate system reliability estimate due to limited data. To overcome this issue, a fuzzy dynamic fault tree model is put forth to assess system reliability. To obtain the membership function of the fuzzy probability for the top event of the studied fault trees, the extension principle is employed to calculate the associated membership function via a pair of parametric programming problems. Finally, a case study is presented to demonstrate the application of the proposed approach for the hydraulic system of a CNC machining centre.

LI YF, HUANG HZ, LIU Y, XIAO N, LI H. **Nowa metoda analizy drzewa uszkodzeń: Rozmyta analiza dynamicznego drzewa uszkodzeń**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 208-214.

Analiza drzewa uszkodzeń (FTA) znajduje szerokie zastosowanie jako narzędzie oceny niezawodności dużych i złożonych systemów inżynierskich. Tradycyjna metoda analizy drzewa uszkodzeń z bramkami logicznymi typu AND, OR, k-z-n, itd. pozwala na sprawne konstruowanie modeli analitycznych reprezentujących kombinacje uszkodzeń elementarnych składowych systemu, które prowadzą do awarii systemu jako całości. Jednakże ograniczone możliwości modelowania jakie daje ta metoda mogą prowadzić do trudności przy modelowaniu systemów dynamicznych posiadających złożone charakterystyki dynamiczne, takie jak zależność sekwencyjna czy zależność funkcjonalna. Analiza dynamicznych drzew uszkodzeń (DFTA) oparta na metodzie Markowa stanowi rozszerzenie tradycyjnej FTA. Wprowadza ona dodatkowe bramki, pozwalając na modelowanie wspomnianych wyżej złożonych interakcji między zdarzeniami. W wielu okolicznościach, ograniczone dane nie pozwalają na otrzymanie dokładnej oceny niezawodności systemu. By rozwiązać ten problem, zaproponowano zastosowanie rozmytego modelu dynamicznego drzewa uszkodzeń do oceny niezawodności systemu. Aby otrzymać funkcję przynależności rozmytego prawdopodobieństwa wystąpienia zdarzenia szczytowego badanego drzewa uszkodzeń, obliczono, na podstawie pary problemów programowania parametrycznego, skojarzoną funkcję przynależności wykorzystując zasadę rozszerzenia. Na zakończenie przedstawiono studium przypadku, w którym proponowane podejście zastosowano do analizy systemu hydraulicznego centrum obróbkowego CNC.

ŚWIĆ A, TARANENKO W. **Adaptive control of machining accuracy of axial-symmetrical low-rigidity parts in elastic-deformable state**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 216-221.

The authors developed a method of correction consisting in the introduction, in the control system, of an additional positive feedback relative to the force of milling. Adaptive control was applied for axial feed and for additional force actions causing the elastic-deformable state, which permits the elimination of static errors of control effects and interference in the control of quality parameters.

ŚWIĆ A, TARANENKO W. **Sterowanie adaptacyjne dokładnością obróbki części osiowo-symetrycznych o małej sztywności w stanie sprężyście-odkształcalnym**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 216-221.

Opracowano sposób korekty ustawienia układu technologicznego obrabiarki polegający na wprowadzeniu do układu sterowania dodatkowego dodatniego sprzężenia zwrotnego względem siły skrawania. Zastosowane sterowanie adaptacyjne posuwem wzdłużnym oraz dodatkowymi oddziaływaniami siłowymi, wywołującymi stan sprężyście-odkształcalny, umożliwia wyeliminowanie błędów statycznych oddziaływania sterującego oraz zakłócającego przy sterowaniu parametrami dokładności obróbki części o małej sztywności w stanie sprężyście-odkształcalnym.

TOMPOROWSKI A, OPIELAK M. **Structural features versus multi-hole grinding efficiency**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 222-227.

The purpose of this paper has been to describe the influence of design features and properties of disintegrated grain biomaterials on the dynamics of the process efficiency. It has been determined that the basis for improved functionality of the grinding machine is the analysis of the influence of the effective component of gravitational force and angle of repose of the grain biomass on the dynamics of changes in the efficiency of the five-disc, multi-hole RWT-5KZ grinder. Reasonable efficiency may be obtained by means of purposeful control of cross-sections and volumes of grinding holes, i.e. design features of multi-hole multi-disc unit. The actions will however bring the planned benefits only when a mathematical description is developed for the flow of the disintegrated grains (biomass grains) through the working space of the multi-disc grinder as a resultant variable of the design and operation of the working unit. The search for design solutions of the units that grind grain, leading to efficient processing justify the research into the improvement of the theory and design of grinding machines.

TOMPOROWSKI A, OPIELAK M. **Konstrukcyjne kształtowanie wydajności rozdrabniania wielootworowego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 222-227.

W pracy podjęto próbę opisu wpływu cech konstrukcyjnych oraz właściwości rozdrabnianych biomateriałów ziarnistych na dynamikę wydajności procesu. Uznano, że podstawą do poprawy funkcjonalności działania maszyny rozdrabniającej jest przeprowadzenie rozpoznania wpływu efektywnej składowej siły grawitacji i kąta usypu biomasy ziarnistej, na dynamikę zmian wydajności pięciotarczowego, wielootworowego rozdrabniacza RWT-5KZ. Racjonalna wydajność może być osiągnięta, między innymi, na drodze celowego sterowania przekrojami i objętościami otworów uczestniczących w rozdrabnianiu, czyli cechami konstrukcyjnymi wielotarczowego zespołu wielootworowego. Aby jednak działania te przyniosły planowane korzyści, konieczne staje się opracowanie opisu matematycznego przepływu rozdrabnianych ziaren zbóż (ziaren biomasy) przez przestrzeń roboczą rozdrabniacza wielotarczowego, jako zmiennej wynikowej konstrukcji i działania zespołu roboczego. Poszukiwania rozwiązań konstrukcyjnych zespołów rozdrabniających ziarna zbóż, prowadzące do wydajnych procesów przetwórczych, uzasadniają podjęcie badań nad doskonaleniem teorii i konstrukcji rozdrabniaczy.

ORKISZ M, ŚWIĘCH Ł, ZACHARZEWSKI J. **Fatigue tests of motor glider wing's composite spar**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 228-232.

The paper presents experimental and numerical investigation of wing's spar. Part of the spar was subjected to one step fatigue test, covering 10,000 load cycles corresponding to the oscillations of the load factor from nz min = − 3.7 to nz max = 5.7. Such test is proposed as an alternative to the full loading spectrum tests. During the experiment 3D scan was used to rapid inspection of sensitive structure's areas. Application of optical strain gauges based on a fiber Bragg's grating allowed to observe the phenomenon of local, periodical strengthening of the structure.

ORKISZ M, ŚWIĘCH Ł, ZACHARZEWSKI J. **Badania zmęczeniowe kompozytowegodźwigara skrzydła motoszybowca**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 228-232.

W pracy przedstawiono badania eksperymentalne i numeryczne dźwigara skrzydła. Fragment dźwigara poddano jednostopniowemu testowi zmęczeniowemu, obejmującemu 10.000 cykli obciążeń odpowiadających oscylacjom współczynnika obciążeń od nz min= −3.7 do nz max= 5.7. Test taki proponowany jest jako alternatywa dla próby z zastosowaniem pełnego spektrum obciążeń. W trakcie badań wykorzystano skanowanie przestrzenne, jako propozycję szybkiej metody inspekcji newralgicznych obszarów konstrukcji. Zastosowanie do pomiaru odkształceń systemu światłowodowych czujników tensometrycznych opartych na siatce Bragga'a pozwoliło na zaobserwowanie zjawiska lokalnego, okresowego umacniania struktury.

JURCA V, ALES Z. **Maintenance management systems in agricultural companies in the Czech Republic**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 233-238.

The purpose of this paper is to describe different maintenance management systems being used by Czech companies with a view to the agro-industrial complex in recent years. The paper focuses on maintenance management systems supported by information system (IS) and their implementation in companies. The use of IS helps to create a variety of scenarios with an emphasis on the need for analytical instruments processing data integration in order to improve the efficiency of the maintenance system and identify its weaknesses. While the maintenance management information system (MMIS) is widespread within the agro-industrial scope in the manufacturing industry, in primary agricultural production and agricultural machinery servicing, it has yet to be utilized. The paper also describes some possible applications of MMIS in the sphere of the agro-industrial complex, the differences between their applications and the potential use of MMIS in the case of fully or partially outsourced agricultural equipment maintenance.

JURCA V, ALES Z. **Systemy Zarządzania Utrzymaniem Ruchu w przedsiębiorstwach rolnych w Republice Czeskiej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 233-238.

Przedstawiona praca ma na celu opis różnych systemów zarządzania utrzymaniem ruchu stosowanych w ostatnich latach przez czeskie przedsiębiorstwa ze szczególnym uwzględnieniem zakładów rolno-przemysłowych. W pracy główną uwagę skupiono na systemach zarządzania utrzymaniem ruchu wspieranych przez systemy informatyczne (IS) oraz na problemie ich wdrażania w przedsiębiorstwach. Użycie systemu informatycznego pomaga w tworzeniu różnych scenariuszy, przy czym nacisk kładzie się na zapotrzebowanie na instrumenty analityczne służące do integracji przetwarzanych danych w celu polepszenia wydajności systemu utrzymania ruchu oraz rozpoznawania jego słabych punktów. Podczas gdy użycie systemów informatycznych wspomagających zarządzanie utrzymaniem ruchu (ang. maintenance management information system, MMIS) jest powszechne w zakładach rolno-przemysłowych przemysłu wytwórczego, podstawowa produkcja rolna oraz obsługa maszyn rolniczych czekają jeszcze na ich wdrożenie. W artykule opisano także możliwe zastosowania MMIS w zakładach rolno-spożywczych oraz różnice w aplikacji i potencjalnym wykorzystaniu MMIS w przypadku prac w zakresie utrzymania ruchu maszyn i urządzeń rolniczych całkowicie lub częściowo zleconych firmom zewnętrznym.

KRUCZYŃSKI S. W . **Maintenance of three way catalytic converter – thermal deactivation**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 239-242.

The present paper gives a description of deactivation processes of three way exhaust gas catalytic converters, with special consideration of thermal deactivation. Test results of new (or "fresh") and aged reactor during five-hours thermal test are presented. Conversion of carbon monoxide, hydrocarbons and nitric oxides is evaluated. Ignition temperatures of catalyst of individual noxious matters are determined, as well for new as for aged catalyst. Results of conversion measurements are correlated with results of physical and chemical tests of catalyst structure changes during ageing test.

KRUCZYŃSKI S. W . **Eksploatacja trójfunkcyjnych reaktorów katalitycznych – dezaktywacja termiczna**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 239-242.

W pracy opisano procesy dezaktywacji trójfunkcyjnych reaktorów katalitycznych spalin ze szczególnym uwzględnieniem dezaktywacji termicznej. Przedstawiono wyniki badań rektora nowego i reaktora starzonego w pięciogodzinnym teście termicznym. Oceniono konwersję tlenku węgla, węglowodorów i tlenków azotu. Wyznaczono wartości temperatury zapłonu katalizatora poszczególnych substancji szkodliwych dla kata-lizatora nowego i starzonego. Wyniki pomiarów konwersji skorelowano z wynikami badań fizyko-chemicznych zmian struktury katalizatora podczas testu starzeniowego.

ZARINCHANG A, FAGHIH N, ZARINCHANG J. **An application of Genetic Algorithm toward solving the reliability problem of multiobjective series-parallel systems.** Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 243-248.

Since developing an appropriate solution for reliability optimization problem with mathematical programming methods has been considered as difficult techniques, the heuristic approaches increasingly has been applied. Multiobjectve Genetic Algorithm (MGA) has been among heuristic methods that was developed to find solutions for series-parallel systems to obtain maximum reliability, and minimum cost and weight at the system level. These are very common problems in engineering design such as mechanical and electrical systems. It has been shown that the Multiobjectve Genetic Algorithm offers proper results to these problems while it respects to the several objective functions such as reliability, cost and weight. This paper presents the combination of probabilistic search, and one of the decision making methods called Technique for

ZARINCHANG A, FAGHIH N, ZARINCHANG J. **Zastosowanie algorytmu genetycznego do rozwiązywania zadań niezawodnościowych dotyczących wielokryterialnychsystemów szeregowo-równoległych.** Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 243-248.

Ponieważ znalezienie odpowiedniego rozwiązania zadania optymalizacji niezawodnościowej przy wykorzystaniu metod programowania matematycznego uznaje się za trudne, coraz częściej stosuje się do tego celu metody heurystyczne. Algorytm genetyczny do optymalizacji wielokryterialnej (Multiobjective Genetic Algorithm, MGA) jest jedną z metod heurystycznych, stworzoną w celu znajdowania rozwiązań dla systemów szeregowo-równoległych, pozwalającą na uzyskanie maksymalnej niezawodności oraz minimalnych kosztów i ciężaru na poziomie systemu. Zadania takie występują powszechnie w dziedzinie projektowania i konstrukcji systemów mechanicznych i elektrycznych. Wykazano, że MGA pozwala uzyskać odpowiednie rozwiązania tego typu zadań uwzględniając przy tym funkcje celu, takie jak niezawodność, koszty i ciężar.

Order Preference by Similarity to Ideal Solution (TOPSIS). The Multiobjectve Genetic Algorithm, allows us to achieve a proper design solution while it saves a considerable time compared with some other approaches. At the same time as the reliability, cost and weight were chosen as objective functions, the results obtained by this method showed an overall improvement in comparison to the existing GA method considering cost and weight as constraints.

W niniejszej pracy przedstawiono połączenie metody wyszukiwania probabilistycznego oraz jednej z metod rozwiązywania problemów decyzyjnych o nazwie TOPSIS (Technique for Order Preference by Similarity to Ideal Solution). MGA pozwala uzyskać odpowiednie rozwiązania konstrukcyjne dając przy tym znaczną oszczędność czasu w porównaniu z niektórymi innymi metodami. Jednocześnie potraktowanie kosztów i ciężaru jako funkcji celu daje lepsze wyniki w porównaniu do metody wykorzystującej algorytm genetyczny, w której koszty i ciężar rozpatrywane są jako ograniczenia.

NIEWCZAS A M, PIENIAK D, OGRODNIK P. **Reliability analysis of strength of dental composites subjected to different photopolymerization procedures**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 249-255.

The aim of this study was evaluation of chosen photopolymerization procedures on strength and reliability of dental composites based on siloranes and composites based on methacrylate compounds in 3-points bending test conditions. The following composites were tested: Filtek Siloran (FSi), Gradia Direct Anterior (GDA), Gradia Direct Posterior (GDP), Herculite XRV (H). Photopolymerization was conducted by means of two types of light: LED lamp and halogen lamp. Exposure times of 40 and 60 seconds were applied. For the strength studies a three-point bending test was used (TFS). Twenty rectangular beam-shaped samples (N=20) from each material were prepared for the studies. For each studied case an average value and standard deviation were determined. To assess significance of differences a variation analysis was performed. Then, the results from each specimen were approximated by two-parameter Weibull distribution. Distribution scale parameter was calculated (as a characteristic strength) and shape parameter (as a material reliability index). It has been demonstrated that in 3-point bending test conditions in case of silorane-based composite the type of lamp has no impact on the strength, however it can improve its reliability. In case of conventional methacrylate-based materials application of LED lamp instead of halogen lamp reduces material strength, but increases its reliability. Additionally, it has been shown that the extension of exposure time – in case of FSi material and halogen lamp, increases material strength, however it has no impact on reliability of the material.

NIEWCZAS A M, PIENIAK D, OGRODNIK P. **Analiza niezawodnościowa wytrzymałości kompozytów stomatologicznych poddanych zróżnicowanym-procedurom fotopolimeryzacji**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 249-255.

Celem badań była ocena wpływu wybranych procedur fotopolimeryzacji na wytrzymałość i niezawodność kompozytów stomatologicznych opartych na siloranach oraz kompozytów opartych na związkach metakrylanowych w warunkach testu na 3-punktowe zginanie. Badano kompozyty o nazwach handlowych: Filtek Siloran (FSi), Gradia Direct Anterior (GDA), Gradia Direct Posterior (GDP), Herculite XRV (H). Zastosowano fotopolimeryzację dwoma rodzajami światła: lampą diodową oraz lampą halogenową. Przyjęto czas naświetlania 40 sek. oraz 60 sek. Do badań wytrzymałości został zastosowany test na zginanie trójpunktowe (TFS). Przygotowano próbki do badań w formie belek prostopadłościennych o liczności N = 20 z każdego materiału. Wyznaczono wartość średnią i odchylenie standardowe dla każdego badanego wariantu. Do oceny istotności różnic przeprowadzono analizę wariancji. Następnie wyniki każdej próby aproksymowano dwuparametrowym rozkładem Weibull'a. Obliczono parametr skali rozkładu (jako wytrzymałość charakterystyczną) oraz parametr kształtu (jako wskaźnik niezawodności materiału). Wykazano, że w warunkach testu na 3-punktowe zginanie rodzaj lampy nie ma wpływu na wytrzymałość w przypadku kompozytu opartego na siloranach, natomiast umożliwia poprawę jego niezawodności. W przypadku konwencjonalnych materiałów opartych na metakrylanach zastosowanie lampy diodowej w miejsce lampy halogenowej obniża wytrzymałość materiału, jednak zwiększa jego niezawodność. Ponadto wykazano, że zwiększenie czasu naświetlania – w przypadku materiału FSi i lampy halogenowej zwiększa jego wytrzymałość, natomiast nie ma wpływu na niezawodność. W pozostałych przypadkach wytrzymałość na ogół pozostaje na stałym poziomie lecz zwiększa się niezawodność materiału.

ZHANG T, CHENG Z, LIU YJ, GUO B. **Maintenance scheduling for multi-unit system: a stochastic Petri-net and genetic algorithm based approach**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 256-264.

Frequent maintenance activities would cause low system availability and require large sums of money. For a multi-unit system, maintenance activities of some units can be combined together to reduce the total maintenance possession time and cost. Therefore, an optimized timetable of the maintenance activities is needed to be planned. Considering the uncertainties in both the deterioration and maintenance process of the units in a system, this paper advances a stochastic Petri-net based simulation optimization model for maintenance scheduling. The genetic algorithm is used to get the solution of the timetable of the maintenance activity schedule such that the overall cost is minimized in a planning horizon taking into account total maintenance possession time, unit condition, life cycle loss and solution feasibility. Some techniques used to reduce the computational effort required to perform the analysis are also described. A case study is given in the end.

ZHANG T, CHENG Z, LIU YJ, GUO B. **Ustalanie harmonogramu obsługi dla systemu wieloelementowego: podejście oparte na stochastycznych sieciach Petriego oraz algorytmie genetycznym**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 256-264.

Częste czynności obsługowe prowadzą do niskiej gotowości systemu oraz wymagają dużych nakładów pieniężnych. W systemie wieloelementowym całkowity czas i koszt obsługi można obniżać łącząc ze sobą czynności obsługowe niektórych elementów. Dlatego też konieczne jest planowanie zoptymalizowanego harmonogramu czynności obsługowych. W artykule zaproponowano model symulacyjny optymalizacji harmonogramu obsługi oparty na stochastycznych sieciach Petriego uwzględniający niepewność zarówno procesu deterioracji jak i procesu obsługi elementów systemu. Algorytm genetyczny wykorzystano do opracowania terminarza czynności obsługowych, który pozwalałby na minimalizację kosztów całkowitych w przyjętym horyzoncie planowania przy uwzględnieniu całkowitego czasu obsługi, stanu elementów, strat wynikających z cyklu życia oraz wykonalności rozwiązania. Ponadto opisano techniki zastosowane w celu zmniejszenia wysiłku obliczeniowego potrzebnego do wykonania analizy. W końcowej części pracy przedstawiono studium przypadku.

Article citation info:
VALIŠ D, VINTR Z, MALACH J. Selected aspects of physical structures vulnerability– state-of-the-art. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 189-194.

David VALIŠ
Zdenek VINTR
Jindrich MALACH

SELECTED ASPECTS OF PHYSICAL STRUCTURES VULNERABILITY – STATE-OF-THE-ART

WYBRANE ZAGADNIENIA DOTYCZĄCE NARAŻENIA OBIEKTÓW FIZYCZNYCH NA ATAK – STAN WIEDZY

The paper is to deal with the selected aspects of structures vulnerability in terms of the physical protection. It is going to specify possible risks following from a terrorist attack, define a potential target and its characteristics, describe the resistance of an object to an attack, and determine the ways to reduce the probability of reaching a terrorist target, or increase object resilience. The results we are going to introduce reflect current knowledge in the area of physical protection.

Keywords: *Physical protection, target, vulnerability, risk.*

W artykule omówiono wybrane aspekty narażenia obiektów fizycznych na atak w kontekście ochrony fizycznej. Określono możliwe zagrożenia wynikające z ataku terrorystycznego, zdefiniowano potencjalne cele ataku i podano ich charakterystyki, opisano zjawisko wytrzymałości obiektu na atak, oraz ustalono sposoby zmniejszania prawdopodobieństwa dotarcia do celu ataku terrorystycznego lub zwiększania odporności obiektu. Przedstawione wyniki odzwierciedlają bieżący stan wiedzy w dziedzinie ochrony fizycznej.

Słowa kluczowe: *Ochrona fizyczna, cel ataku, narażenie na atak, zagrożenie.*

1. Introduction

Modern history is rich with examples of various terrorist attacks against structures, transportation systems, etc. worldwide. In the aftermath of the September 11th tragedies, the vulnerability of the all infrastructure to terrorist attack has gained national attention. In light of this vulnerability, various governmental agencies are looking into ways to improve the design of structures to better withstand extreme loadings. Tens of per cent of the homeland security outlays are devoted by countries to making potential targets less vulnerable to potential terrorist attacks. This is to protect what we usually call "Critical Infrastructure", "Key Asset" and/or "Key Resources". The concern of this paper is on the so called "passive defence" which is one part of several others like safe buildings, drinking water protection, Rapid Risk Assessment, etc. [14] We would like to discuss issues related to defining threat, describe vulnerability of possible targets and assess the possibilities of protection efficiency.

In terms of possible attack unpredictability is a key characteristics terrorist activity, for two main reasons [13]:

Terrorist have many more categories of legitimate targets, as well as worldwide scope, compared to traditional security concerns (which used to have the comparable luxury of protecting obvious military assets, or home territory).

Terrorist attack can have different objectives like to harm people, to damage infrastructure, to cause panic, etc.

Although such objectives may often overlap, these varying objectives lead to varying types or location targets. However, we have to keep in mind that detection and prevention must always remain the first line of defence [18].

Risk – definition and levels

We understand risk here as it is defined in [5]. The risk event is that the attack on a target with various outcomes. We have also to speak about the "acceptable risk level" based on the target selection and possible threat / consequences description. Taking into account work done by Stewart [19] there is after massive review decided to define following. The "global" consensus or generic quantitative safety goals obtained for involuntary fatality risk to an individual are thus:

- Annual fatality risk higher than 1×10^{-3} are deemed unacceptably high;
- Annual fatality risk in the range 1×10^{-3} to 1×10^{-6} are generally acceptable if the benefits outweigh the risk to provide an economic or social justification of the risk
- Annual fatality risk smaller than 1×10^{-6} are deemed as negligible and further regulation is not warranted.

The individual annual fatality risk can be therefore expressed:

$$\Pr(L) = \sum_{H}\sum_{DS}\sum_{L} \Pr(H)\Pr(DS|H)\Pr(L|DS) \quad (1)$$

where Pr(H) is the annual probability of hazard occurrence per item of infrastructure, Pr(DS | H) is the conditional probability of a damage state (e.g. safety hazard) given occurrence of the hazard and Pr(L | DS) is the conditional probability of a loss (e.g., damage costs, fatalities) given occurrence of the damage state.

We can also estimate expected cost spent on risk reduction per life saved. Protective measures will reduce fatality risks, with a reduction in expected fatalities of p_{attack} Pr(L | H)RN/100 where N is the number of people exposed to the hazard. It follows that the expected cost spent on risk reduction per statistical life saved (E_{LS}) is:

$$E_{LS} = \frac{100C_R}{p_{atteck}\Pr(L|H)RN} \quad (2)$$

where p_{attack} Pr(L | H) is the baseline individual annual fatality risk assuming no protective measures, C_R is the annual cost spent on protective measures for the item of infrastructure and R is the percentage risk reduction as a result of protective measures. We do not fully agree with this formula since risk reduction R plays role also in the Pr(L | H) as well as in the p_{attack}. Therefore the value R shall not be explicitly mentioned in the equation (2). On the other hand the percentage risk reduction modifies significantly the cost spent.

For illustration Figure 1 shows that a 95% reduction in risks results in annual fatality risks at least an order of magnitude lower than the quantitative safety goal [19].

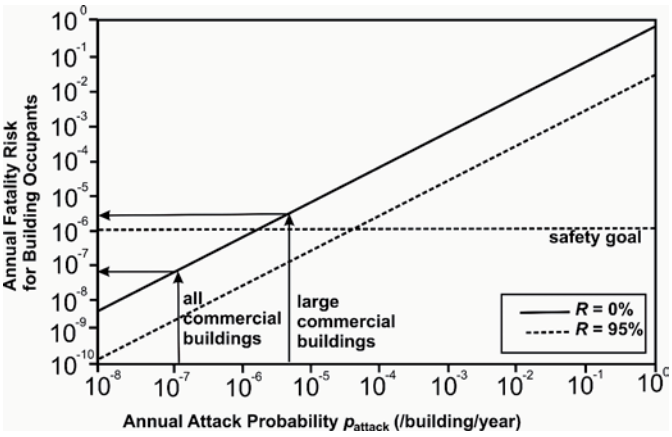


Figure 1. Individual annual fatality for building occupants risk showing quantitative safety goal of 1×10^{-6} fatalities/year

Since Pr(L) = $0.5p_{attack}$, equation (3) can then be re-expressed as:

$$E_{LS} = \frac{100C_R}{0.5p_{atteck}RN} \quad (3)$$

2. Discussing threat

Before developing a plan to enhance building security, a risk manager shall first gain an understanding of potential threats acting against a structure/building as well as understanding of terrorists' goals and tactics. There are numerous combinations of explosive devices, cutting devices, impact vehicles and specific attack location to consider. It is impossible to design and protect all structures/buildings to resist all possible combinations of terrorist attacks that may occur. Simply said there are too many possible combinations. Therefore, a process is proposed to determine the most likely tactics from the terrorists' perspective and reduce number of possible combinations of attack that need to be considered for the purpose of design [20]. Various threat characteristics (against which a physical protection is prepared) are defined by a state authority. This characteristic in in Czech Republic called as "Design Basis Threat". Such characteristic includes numbers, equipment, knowledge and skills, intentions, strategy, etc.

Terrorist goal will vary depending upon the specific interest of different domestic or international organisation. However, goals which are most often encountered include making a high visibility statement through media coverage, obtaining publicity for their cause, destroying landmark or critical assets, exerting political pressure, advancing religious imperative, seeking vengeance, creating public fear and panic, maximizing casualties, disrupting the economy, and interrupting main or emergency transportation routes. Based on statement in [13] it is highly improbable that near-optimal stations were targeted by chance.

Due primarily to the amount of preparation time required, terrorists are not typically/likely to use conventional civilian or military structure demolition tactics. These techniques require specialized skills and considerably more time on target in order to achieve the precision demolition effects. Terrorist generally seek simple, rapid, and flexible plans; therefore, the tactics that they are most likely to use include making bomb threats to disrupt the structures, transportation systems; employing hand-placed explosives or cutting devices in critical hidden areas of a building – if possible. Than using fragments to target vehicles during times of high traffic density; targeting multiple structures to disrupt the infrastructure systems and reduce emergency responder effectiveness; cutting critical utilities running on bridges; using collapsed span as an obstacle to block a critical waterway or destroy a nearby facility; making use of trucks, ships, trains, or planes as impact or explosive vehicles.

Terrorist plans include at least following characteristics:
- Realistic;
- Coordinated;
- Cohesive;
- Simple;
- Creative;
- Flexible; and
- Secretive.

In deciding which strategy to use, terrorists will probably use comparison criteria such as:
- Having a high probability of success;
- Being realistic with easily obtainable materials;
- Occurring quickly to minimise placement and priming time;
- Utilising secrecy and surprise to minimise chances of being detected;
- Being simple and easy to coordinate;
- Providing flexibility to change the plan;
- Having a large impact and magnitude to maximize damage, casualties, and public panic, etc.

Course of action involving vehicles generally perform well when considering all these criteria, and therefore tend to be the "most likely" courses used. Scenarios involving hand-placed explosives limit the charge weight that can be used, and thus reduce the overall impact and chances of success unless they can achieve sufficient time on target. Most of the cases where smaller charges would be very effective involve locations that can often be difficult to access, thereby reducing their speed, simplicity, and flexibility. Additionally, the threat of hand-placed precision demolitions can be readily countered with physical security and detection measures on a structure/building.

Threat therefore modifies intended levels of physical protection as well as risk profile in terms of possible attack. This all might be described by common mathematical characteristics.

Threat levels

Protective measures could be divided into threat levels for implementation as the appropriate level is reached, as demonstrated in table 1. Each country has its own "Common threat levels" e.g. according to the legislative for the critical infrastructure protection. Countermeasures are therefore modified by such possible threat levels. In Table 1 there is modified scale in accordance with [20].

Table 1. Scale levels of threat levels and measures

Threat level to a structure/ building	Additional security measures
Severe	Restricted access with guards, barriers, and vehicle searches. All other measures listed bellow.
High	Increased frequency of patrols and checks. Conduct unscheduled exercise of emergency response plans. Postpone nonessential maintenance. Coordinate with national security corps or law enforcement for possible closure and vehicle searches once severe level is reached. All other measures listed bellow.
Elevated	Implement regularly scheduled police patrols. All other measures listed bellow.
Guarded	Review and update emergency response procedures. Increase frequency of periodic checks of cameras, fences, etc. All other measures listed bellow.
Low	Monitor security systems in place (including periodic checks). Disseminate threat information to personnel. Regular refinement and exercising of emergency operations plan. Emergency responder training. Continually updating threat and vulnerability assessment.

Each of the security measures stated above is worth respective amount of money.

3. Defining vulnerability

Department of Homeland Security (DHS) in the USA defines vulnerability as "physical feature or operational attribute that renders an entity, asset, system, network, or geographical area to exploitation or susceptible to a given hazard" (2010) [3]. The key of assessing vulnerability properly is in the last phrase of that definition. Although vulnerability assessments can be standalone documents, vulnerability is best understood within a risk context, specifically the interaction between the threat and the consequence. This interaction is the reason that vulnerability V is sometimes defined as the probability of success (of an attack) P_S given an attack A or probability of the consequence occurring given an event. Mathematical expression is than:

$$V = P_S(\mathrm{A}) \tag{4}$$

In either case vulnerability is the collective influence of physical features or operations that reduce the effectiveness (alternatively success) of the adversary's attack or that make the target better able to sustain the attack. Analysis is highly dependent, therefore, on the method of attack and strength of the attack expected. A building's vulnerability to an improvised explosive device (IED) will differ from the vulnerability to a vehicle-born IED (VBIED), for example, depending on the assumption in the definition of those attacks, such as amount or type of explosives, entry points, and stand-off distance. Even within the category of VBIED, vulnerability will differ based on terrorist tactics, such as leaving the vehicle on the street adjacent to the building or ramming the vehicle into a building or its defensive perimeter. The more specific the context, the more accurate the vulnerability assessment for particular target can be.

For security risk, vulnerability is also influenced by the terrorist adversary. Terrorist groups have different levels of competence and expertise. This can affect not only target selection but also their knowledge of countermeasures and their determination to overcome those countermeasures through technology or effort. These aspects of the threat can influence judgements of degree of accessibility or strength of countermeasures (this states actually one of very hardly measured characteristics physical protection system which is the deterrence. But for well-equipped and trained terrorist discourage is very low however). Opportunity to attack, in other words, reflects the interaction of threat and vulnerability; the characteristics of potential attackers help provide further context for high-quality vulnerability assessments. With all of these variables, it is easy to see why some argue that vulnerability is not a static characteristic but very dynamic state and, in the extreme, a combination of the various states of all the aspects of the asset, facility, or system, which is in constant flux. There always still need for simple way to generate a repeatable and comparable vulnerability level that is useful for the user (government for instance) in the infrastructure protection.

For this reason we have to accept kind of conceptual approach to vulnerability assessment of structures as mention for instance in [3]:

1. Characteristics of the asset itself;
2. The protective measures that prevent access for attack;
3. Access allowed to outsiders and insiders;
4. The functional dependencies on internal and external entries;
5. Generating scenarios
6. Attack methods filtering
7. Event/fault tree analysis (recognisability, countermeasures effectiveness, robustness/resistance)
8. Combining the components

If we speak about vulnerability we cannot forget also to emphasise the structural robustness. It might be expressed by "Protection categories" as said in [21], "Robustness Index" as mentioned in [4] and has several degrees on scale – usually 1-10. Some retrofit recommendations for increasing the structure robustness are for instance listed in [2]. Considering the further statements in [4] there are three most important structural properties which increase a structure's/ building's ability to survive catastrophic overload or damage:

- Structural redundancy (A structure that will perform well in catastrophic situation will permit gravity loads that must be supported during the event to be carried to the foundations using multiple load paths).
- Fireproofing toughness (A structure's ability to resist fire is an important contribution to its robustness, since fire is often a part of catastrophic event).
- Connection robustness (Structural connections are very important and are critical in holding a building together during the large movements that occur in a fire or other catastrophic event).

There are several ways for assessing the severity of possible terrorist attack. Many of them are based on conventional standards like [5, 7, 8, 9, 10, 11, 12]. In [16] there are also mentioned some possible tools for risk assessment either software-based (e.g. RAMPART – Risk Assessment Method-Property Analysis and Ranking Tool; CONTAMW – software for vulnerability assessment; HVAC – software for heating, ventilation, and air condition in buildings assessment) or classical (standards and books). One interesting procedure is mentioned in [1] and is based on risk approach. In defining the problem and deciding an appropriate scope given the time frame and resources there are four critical targets identified for the risk based methodology.

1. Identifying potential targets of attack, methods of attack, and courses of actions.
2. Deciding which possibilities merit deeper scrutiny.
3. Identifying a mathematical way to represent intelligence data in the model.
1. Integrating the first three tasks into a framework that yields output useful in fulfilling the objectives.

The methodology proposed – called "Risk Filtering and Ranking Method (RFRM) addresses what can go wrong, what can be done and deciding which possibilities merit deeper scrutiny. The methodology uses RFRM to identify the most critical contributors to the risk associated with a potential terrorist event to focus the rest of the analysis on. RFRM considers both quantitative factors, such as severity as measured by number of deaths or injuries, and qualitative factors, such type of attack. Since the number of components under consideration often can be large, RFRM is very useful in filtering and prioritizing scenarios.

3.1. Vulnerability characteristics

If we are to talk about building's vulnerability, we have to bear in mind the following facts [17].

- The number of potential terrorist targets is essentially infinite (Terrorists seek to kill people and/or destroy property in pursuit of political goal).
- The number of terrorists appears to be exceedingly small and their efforts and competence rather limited (In 2002 an intelligence report were asserting that the number of trained al-Qaeda operatives in USA was between 2.000 and 5.000).
- In many cases the target selection is effectively a random process (This process, together with other internal motivating mechanisms stressing group cohesion and camaraderie more than grand planning, effectively make terrorist target selection something like random process. Efforts to determine terrorist "intent" in advance become, then, highly problematic.
- The probability that any specific target will be attacked is extremely small in almost all cases (Despite the attention in garners, terrorism is rather rare occurrence comprised of incidental, isolated acts of mayhem perpetrated by individuals or small groups, violence that generally does a comparatively limited amount of damage. Even under quite dire scenarios, in country like the USA, the chance an individual target will be hit is vanishingly small).
- If one potential target happens to enjoy a degree of protection, the agile terrorist generally can readily move on to another one (There is also something that might be called "the displacement effect" Terrorists can choose and change their targets depending on local circumstances. There have been instances in Israel in which the suicide bombers, seeing their primary targets, shopping malls, rather well protected, blew themselves up instead on the street).
- To the degree protection measures make one target safer, they make other ones less safe (For example, there is a program to protect bridges in the USA, and a list of something like 200 of the most important bridges had been drawn up. There seems to be no evidence terrorists have any particular desire to blow up a bridge, due in part, perhaps, to the fact that it is an exceedingly difficult task under the best of circumstances, and the number of casualties is likely to be much lower than for many other targets.
- Most targets are "vulnerable" in that it is not very difficult to damage them, but invulnerable in that they can be rebuilt in fairly short order and at tolerable expenses (on the one hand, most, probably almost all, potential terrorist targets are "vulnerable" in the sense that they can be damaged, in many cases badly, even by a simple explosion).
- It is essentially impossible adequately to protect a very wide variety of potential terrorist targets except by completely closing them down (As it happened, the bombs did no damage because they were poorly constructed and did not actually explode, but this fortunate result, of course, stems entirely from terrorist incompetence, not from the protective measures).

As stated in [17] for example, the applications leading to resistance increase are appropriate in case they take a required effect. These are:

- Nuclear and chemical plants and material (there are not large number of nuclear plants, and an adept terrorist attack on them could potentially have devastating consequences. Consequently, they seem to be prime candidates for protection).
- Key infrastructure nodes (unfortunately it is not at all clear that any such nodes exist although they are some in the EU legislative and some also in respective countries legislation.
- Major ports.
- Symbolic structures – potential targets (religious, historical, etc.).
- Others.

Based on the principles presented in [15] we can understand nine criteria or variables selected for constructions vulnerability assessment like:

- Visibility level of the site ("0" – invisible up to "5" – very high visibility).
- Criticality of the site to its jurisdiction (e.g., city or town "0" no usefulness up to "5" – critical).
- Impact of the site outside of the jurisdiction ("0" – none up to "5" – very high).
- Accessibility of the site to the public ("0" – restricted up to "5" – unlimited).
- Possible hazard located on the site (like Weapons of Mass Destruction – WMD or CBRNE. "0" – none up to "5" – very high).
- Height of the structure ("0" – underground up to "5" – sky scraper).
- Type of the structure ("0" – underground up to "5" – wood structure)
- Population capacity on the site ("0" – no population up to "5" – more than 50.000 people).
- Potential for collateral mass casualties ("0" – 0-100 people up to "5" – more than 5.000 people).

The total number of points formed by the above nine scale should be got by addition between 0 and 45. Vulnerability categories are broken down into five groups as follows:

- Negligible vulnerability – total ranking score 0-9 points.
- Low vulnerability – total ranking score 10-18 points.
- Medium vulnerability – total ranking score 19-27 points.
- High vulnerability – total ranking score 28-36 points.
- Critical vulnerability – total ranking score 37-45 points.

3.2. Approaches for solving vulnerability

There are more options for solving the vulnerability of structures. One is based on the Israeli experience with thousands of armed attacks and the proposed structure of SEPHRA (SEcurity Protection and Hardening Risk Analysis) which was successfully used worldwide in the last 20 years in numerous projects. Scheme of principles is on the figure 2 bellow.

We would also like to present several proposals for vulnerability minimisation of possible targets [17]:

1. Planning
   a) Updating the emergency operation plans/crisis management plan to include response and recovery to a terrorist threat involving important structures.

b) Communication and coordination with local and state law enforcement agencies to obtain intelligence, training, and technical support.

c) Regular drills, table-top exercises, and full scale simulations to test response procedures, communication, and coordination.

d) Planning additional redundancy in transportation system through alternate routes, traffic management, modified lane usage, etc.

e) Planning for prompt debris removal and repairs to ensure rapid restoration of services and restore public confidence in the structure.

f) Developing a training plan for maintenance personnel to be observant of surroundings and capable of dealing with suspicious objects.

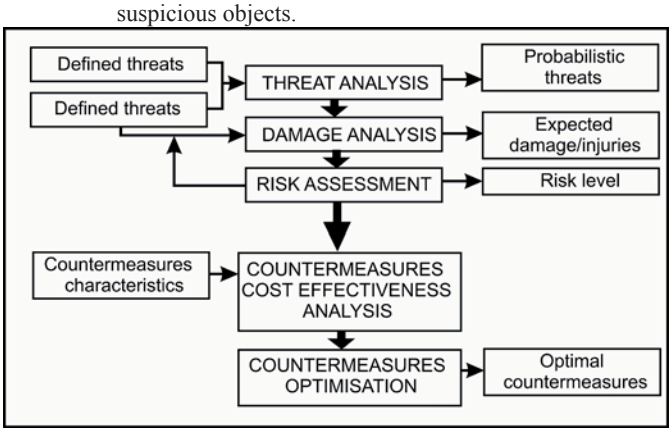


Figure 2. The SEPHRA diagram

2. Information control:
   a) Establish "need-to-know basis" procedures for the release of vulnerabilities, security measures, emergency response plans, or structural details for specific structure.
   b) Review and sanitize websites for potential information which may be beneficial for terrorists.
3. Site layout measures:
   a) Improved lighting with emergency backup, combined with the elimination of hiding spaces which could be used to prepare explosive charges.
   b) Clearing overgrown vegetation to improve lines of sight to critical areas.
   c) Using creative landscaping with regular maintenance to increase vehicular standoff distance to important structural components.
   d) Elimination of access to critical areas such as beneath the deck, maintenance rooms, etc.
   e) Elimination of parking spaces inside or around/beneath the structure.
   f) Providing pass-through gates in concrete median barriers to enable rerouting of traffic and access to emergency vehicles.
   g) Planning redundancy in individual future structures/buildings.
   h) Avoiding architectural features that may magnify blast effects.
4. Access control:
   a) Police patrols, surveillance, and guards.
   b) Keyed or keyless entry systems on access panels, tower entrances, and maintenance areas.
   c) Exterior and interior intrusion detection systems (boundary penetration sensors, volumetric motion sensors, and point sensors, etc.).
   d) Closed circuit television placed where in cannot be easily damaged or avoided, while providing coverage of critical areas to monitor activity, detect suspicious actions, and identify suspects.
   e) Incorporate a higher level of identification procedures and verification of credentials for maintenance personnel.
   f) Deny/limit access to critical structural elements (i.e. providing fencing around important building parts, restricting access to some places of structures, etc.).
   g) Physical barriers to protect gates, towers, piers, etc.
   h) Physical barriers to control access to the structures during credible threat (use conjunction with random vehicle search).
   i) Rapid removal of abandoned vehicles.
   j) No-fly zones around and above critical structures.
   k) Emergency telephones to report incidents or suspicious activity.
   l) Use of an advanced warning system, including warning signs, horns, and popup barricades to restrict access after span failure.
5. Deception measures:
   a) Installing dummy CCTV cameras to augment active cameras when resources are limited.
   b) Parking an abandoned police vehicle nearby.
   c) Posting intrusion detection signs and warnings.
   d) Effectiveness assessment of physical protection systems.

4. Conclusion

This paper is to analyse the present state of the selected aspects of physical protection. It is quite obvious that terrorist attacks can occur at any time and any place. We are not going to tackle the motives for a terrorist act, or different types of attacks which might be performed by an individual (a recent event in Norway) whose motives are quite personal, or by a group (controlled by an organization) the motives of which are religious or political.

The article is focused on two basic aspects related to physical safety. It is a potential target of an attack, the selection of a target versus the target vulnerability. The paper is to assess both possible risks as a consequence of the attack and objects vulnerability. All the current and commonly used approaches listed above are recommended for possible modification.

It is assumed that this work is by no means the end and later it will focus more on assessing the efficiency of physical protection systems and more precise determination of risks resulting from a potential attack. Therefore the authors will concentrate on specifying in a qualitative and quantitative manner the probability of a successful attack, or the probability of an object to resist.

Acknowledgement: *We gratefully acknowledge the Czech Ministry of the Interior supporting us under project "Security Research – Target Identification VG 20112015040 and with the support of the "Project for institutional development of K-202", University of Defence, Brno.+9971988*

References

1. Blais R A, Henry M D, Lilley S R, Pan J A, Grimes M, Haimes Y Y. Risk-based methodology for assessing and managing the severity of terrorist attack. IEEE Systems and Information Engineering Design Symposium, SIEDS '09 , art. no. 5166175, 171-176.
2. Eytan R. Cost effective retrofit of structures against the effect of terrorist attack – the Israeli experience. Proceedings of the Structures Congress and Exposition 2005, 2161-2172.
3. French G S, Gootzit D. Defining and assessing vulnerability of infrastructure to terrorist attack. Vulnerability, Uncertainty, and Risk: Analysis, Modelling, and Management - Proceedings of the ICVRAM 2011 and ISUMA 2011 Conferences, 782-789.
4. Iding R H. A methodology to evaluate robustness in steel buildings – Surviving extreme fires or terrorist attack using a robustness index. Proceedings of the Structure Congress and Exposition 2005, 511-515.
5. ISO/IEC Guide 73:2009 Ed 2.0 Risk management – Vocabulary – Guidelines for use in standards.
6. ISO 31 000:2009 Ed 1.0 Risk management – Principles and guidelines on implementation.
7. ISO 31 010:2009 Ed 1.0 Risk management – Risk Assessment Techniques.
8. ISO 13 824:2009 Ed 1.0 General Principles on Risk Assessment of systems involving structures.
9. ISO 61 508-(1-7):2008 Ed 2.0 Functional Safety of Electrical/Electronic/Programmable Electronic Safety Related Systems.
10. JCSS (Joint Committee on Structural Safety) – Principles, System Representation × Risk Criteria.
11. ECSS (European Cooperation for Space Standardisation) – Q-ST-40-02C Space Product Assurance – Hazard Analysis.
12. MIL-STD 882B Standard Practice for System Safety.
13. Jordán F. Predicting target selection by terrorists: A network analysis of the 2005 London underground attacks. International Journal of Critical Infrastructures 2008; 4 (1-2): 206-214.
14. Kelly T J, Hofacre K C, Derringer T L, Riggs K B, Koglin E N. Testing of safe building detection technologies and other homeland security technologies in EPA´s Environmental Technology Verification (ETV) program. Proceedings of the Air and Waste Management Association's Annual Meeting and Exhibition 2004, 3449-3457.
15. Kemp R L. Assessing the vulnerability of buildings. Journal of Applied Fire Science, 2007; 14 (1): 53-61.
16. Marshall H E, Chapman R E, Leng C J. Risk mitigation plan for optimizing protection of constructed facilities. Cost Engineering 2004; 46 (8): 26-33.
17. Mueller J. Assessing Measures Designed to Protect the Homeland. Policy Studies Journal 2010; 38(1): 1-21.
18. Stewart M G. Cost effectiveness of risk mitigation strategies for protection of buildings against terrorist attack. Journal of Performance of Constructed Facilities 2008; 22(2): 115-120.
19. Stewart M G. Life-Safety risks and optimisation of protective measures against terrorist threats to infrastructure. Structure and Infrastructure Engineering 2011; 7(6):431-440.
20. Williamson E B, Winget D G. Risk management and design of critical bridges for terrorist attacks. Journal of Bridge Engineering 2005; 10(1): 96-106.
21. Zehrt Jr. W H, Acosta P F. Analysis and design of structures to withstand terrorist attack. Proceedings of the Structures Congress and Exposition 2003, 585-592.

Assoc. Prof. David VALIŠ, Ph.D.
Department of Combat and Special Vehicles
Faculty of Military Technologies
University of Defence
Kounicova 65, 662 10 Brno, Czech Republic
E-mail: david.valis@unob.cz

Prof. Zdenek VINTR, Ph.D.
Department of Combat and Special Vehicles
Faculty of Military Technologies
University of Defence
Kounicova 65, 662 10 Brno, Czech Republic
E-mail: zdenek.vintr@unob.cz

Jindrich MALACH, Ph.D.
EBIS, Ltd.
Krizikova 2962/70a, 612 00 Brno, Czech Republic
E-mail: jmalach@ebis.cz

Article citation info:
PANG Y, HUANG HZ, XIAO NC, LIU Y, LI YF. Possibilistic reliability analysis of repairable system with omitted or delayed failure effects. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 195-202.

Yu PANG
Hong-Zhong HUANG
Ning-Cong XIAO
Yu LIU
Yan-Feng LI

POSSIBILISTIC RELIABILITY ANALYSIS OF REPAIRABLE SYSTEM WITH OMITTED OR DELAYED FAILURE EFFECTS

POSYBILISTYCZNA ANALIZA NIEZAWODNOŚCIOWA SYSTEMU NAPRAWIALNEGO Z POMINIĘTYM LUB OPÓŹNIONYM EFEKTEM USZKODZENIA

Within the practical problems in industrial engineering, the failure effect sometimes can be omitted or delayed if it has less effect on the system. In detail, the prominent features of the system can be described as follows: 1) if a repair time is sufficiently short (less than some threshold value) that does not affect the system operation, i.e. the pessimistic effect of system failure could be ignored. The system can be considered as operating during this repair time. It is called the system with repair time omission (failure effect omitted). 2) if a repair time is longer than the given threshold value and the failure effect is finally suffered. Then the system can be considered to remain operating from the initial stage of the repair till the end of the repair threshold. It is called the system with delayed failure effect. Based on the above two characteristics, model for the related repairable system is introduced in this paper. Two scenarios are discussed where the threshold value is regarded as a constant and non-negative random variable, respectively. Reliability indices such as instantaneous possibilistic availability are formulated for the system with failure effect omitted or delayed.

Keywords: *failure effect omitted or delayed, Markov model, repair time omission, instantaneous possibilistic availability.*

Przy rozwiązywaniu problemów praktycznych w inżynierii przemysłowej można czasami pominąć bądź opóźnić efekt uszkodzenia jeśli ma on niewielki wpływ na system. Ściślej, wiodące cechy systemu można opisać w następujący sposób: 1) jeżeli czas naprawy jest wystarczająco krótki (krótszy niż pewna wartość progowa), tak iż nie ma on wpływu na działanie systemu, to można pominąć negatywny efekt uszkodzenia systemu. Przy takim czasie naprawy można uznać że system nie przerwał działania. Nazywa się go wtedy systemem z pominięciem czasu naprawy (pominięty efekt uszkodzenia). 2) Jeżeli czas naprawy jest dłuższy niż dana wartość progowa i efekt uszkodzenia staje się w końcu odczuwalny, to uznajemy, że system pozostawał aktywny od początkowego etapu naprawy aż do momentu, w którym został przekroczony próg czasu naprawy. Nazywa się go wtedy systemem z opóźnionym efektem uszkodzenia. W oparciu o powyższe dwie charakterystyki, wprowadzono w prezentowanej pracy model systemu naprawialnego. Omówiono dwa scenariusze, w których, odpowiednio, przyjęto, że wartość progowa jest wartością stałą lub nieujemną zmienną losową. Sformułowano wskaźniki niezawodnościowe, takie jak posybilistyczna gotowość chwilowa, dla systemu z pominiętym lub opóźnionym efektem uszkodzenia.

Słowa kluczowe*: pominięty lub opóźniony efekt uszkodzenia, model Markowa, pominięcie czasu naprawy, posybilistyczna gotowość chwilowa.*

Notation

t	time scale
T_S	system lifetime
T_R	system repair time
T_{S_i}	system lifetime within i th period
T_{R_i}	system repair time within i th period
E	state space of the original system
W	working state space of the original system
F	failure state space of the original system
τ	threshold value
$\pi_{T_S}(t_S)$	possibility distribution for variable T_S
$\pi_{T_R}(t_R)$	possibility distribution for variable T_R
Π	possibility measure corresponding to possibility distribution π
$\pi_{(T_S,T_R)}(t_S,t_R)$	joint possibility distribution for variables T_S and T_R
$H(\tau)$	probabilistic distribution for non-negative random variable τ
$A_\Pi(t)$	possibilistic availability of the original system at time t
$\tilde{A}_\Pi(t)$	possibilistic availability of the new system at time t
Z_i	system lifetime plus repair time within ith period $(T_{S\_i}+T_{R\_i})$

1. Introduction

Repairable system is defined as a system which, after failing to perform at least one of its required functions, can be returned to performing all of its required functions satisfactorily by any method other than replacement of the entire system [1]. Reliability analysis of repairable system is a momentous branch of system reliability theory, maintaining a high level of reliability is often an essential requisite for repairable systems [11]. Much research has been devoted to analyze system reliability of repairable system and various models have been proposed on repairable systems [12, 17-20].

Traditionally, the models for repairable system in the literature are involved with a basic assumption [12, 17-19]. The system instantaneously falls into the failure state when it is out of work. However, in some practical situations, for instance, if the demands on the system by 'customers' are not too frequent, they more likely miss a small repair time or are at most delayed by a negligibly short time in receiving service. Within such repairable system, the effect of system failure can be neglected if the repair time is sufficiently short (less than some given critical value) or delayed if the repair time exceeds the critical threshold. In other words, if the system is temporarily under repair and has no effect on system operating, the system can be regarded as being operating during such a repair interval (failure effect omitted). If the system is out of work and fails to back operating before the threshold repair time, it can be regarded remain operating during the period of repair until the repair time exceeds the critical value (failure effect delayed).

Recently, several related repairable systems have been studied. Zheng [20] first proposed this omitted failure model with single-unit repairable system in which repair time that is sufficiently short does not result in a system failure. Based on the introduction of the new model for repairable system, researchers [9, 10, 15, 21] analyzed system availability of series-system, parallel system, K-out-of-N: G system and so on. Furthermore, researchers [2, 22] established and analyzed the model for single-unit, and series repairable system in which failure effects could be neglected or delayed.

Among the research for repairable system in which failure effects could be neglected or delayed, probability theory is the most commonly used theory to analyze system reliability indices. In fact, possibility theory has been used increasingly to model epistemic uncertainty in reliability engineering [3, 7, 8, 13, 16]. This type of uncertainty describes subjectivity or lack of information. It is defined as reducible uncertainty and subjective uncertainty, since it can be reduced with increased state-of-knowledge or collection of more data. It exists extensively within the research of reliability, such as poor understanding of initiating events, fault trees, and event trees [14]. The aim of this paper is to theoretically analyze the system possibilistic availability for repairable systems with omitted or delayed failure effects.

In the next section, basic assumptions of the original model and the new model for single-unit repairable system are introduced, and the new model is distinguished from the original model. A comprehensive possibilistic reliability analysis of the new model is presented in Section 3. A numerical example is given to illustrate the results in the subsequent section. Finally, conclusion is given in Section 5.

2. Basic assumptions of the original model and new model for repairable system

In this section, original model and new model for single-unit repairable system are introduced respectively, as well as the difference between them. Firstly, the assumptions of the original model are addressed before presenting the new model [4, 5]:

1) The system is composed of one component and one repair facility. The system is new at the initial time (t=0), and when the component fails, the repair begins immediately.
2) The system has two possible states: up (operating), and down (failed). The repaired component is restored into "as good as new" condition.
3) Assume that the system lifetime T_S follows deflection minor type possibility distribution $\pi_{T_S}(u)$, and the repair time T_R follows deflection minor type possibility distribution $\pi_{T_R}(u)$. All the variables involved are mutually independent.

Suppose that state 0 represents the failed (down) state and state 1 be the operating (up) state, the system state space is denoted as $E=\{0,1\}$. The working state space is $W=\{1\}$ while the failure state space is $F=\{0\}$. Let $X(t)$ denote the stochastic process of the system state at time instant t, thus one has:

$$X(t)=\begin{cases}1, & \text{the system is in the operating state at time } t;\\ 0, & \text{the system is in the failed state at time } t.\end{cases}$$

In fact, $\{X(t), t\geq 0\}$ forms a homogeneous continuous time process in state space E. For such single-unit system, the state of system is just the same with the state of unit. Based on the assumptions of original model, new model that presents a different way of system operation is addressed. The main difference between the original model (single-unit repairable system) and the new model (single-unit repairable system with failure effect omitted or delayed) is that, the new system may still be in the operating state while it turns out to be failed within the original system at the meantime. In details, given a threshold value τ $(\tau\geq 0)$, if the involved repair time is shorter than τ, the new system can be thought of being in the operating state during the repair interval. If the repair time exceeds the threshold value τ, the new system is regarded to be still under the state of operating during the repair interval of $(0,\tau)$ and being in the failed state from the repair time τ till the end of the repair. In other words, for the case that repair finishes before reach the threshold τ, the new system is considered to be under operating state during the repair interval while the original system being in failure state during the same period. Thus, the failure effect has been omitted. For the other case that repair can't finish before the given threshold value τ, the new system is regarded as under the state of operating during the process of repair until the given threshold value τ is exceeded. During the same period, the original system is under the failure state. Thus, the failure effect has been delayed for τ.

As for the same single-unit repairable system that is composed of one component and one repair facility, the new model is described as follows in contrast with the original model:

1) The system is new at the initial time (t=0), and failed unit will receive repair immediately after failure. All the random variables are mutually independent.
2) The system has two possible states: up (operating), and down (failed). The repaired component is restored into "as good as new" condition.
3) The system failure time T_S and the system repair time T_R follow deflection minor type possibility distribution $\pi_{T_S}(t_S)$ and $\pi_{T_R}(t_R)$, respectively as in the original model.
4) The new system is operating if the original system is operating.
5) If the original system fails and the repair time is less than the threshold value τ, then the new system is still operating during the repair time.

6) If the original system fails and it takes longer than τ to finish the repair, then the new system is considered to remain in the operating state during the repair interval of $(0,\tau)$, and being in the failed state after this interval until the repair complete.
7) The threshold value τ can be either a constant or a non-negative random variable. τ follows the distribution of $H(\tau)$ if it is considered as a non-negative random variable.

In particular, when $\tau = 0$, the new system becomes the original system. When $\tau = \infty$, the new system is never down. A possible sequence of system state changes of the original system and the new system is shown in Fig.1.

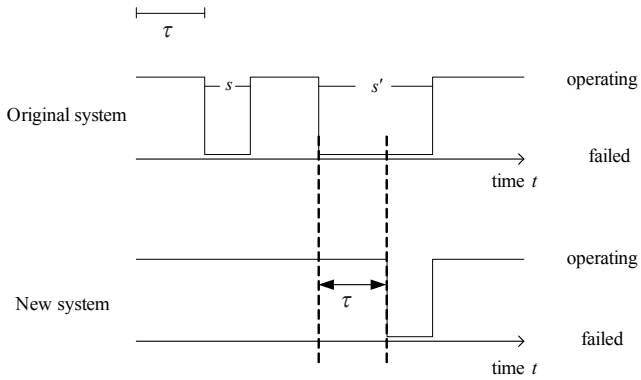


Fig.1. Difference between the original system and the new system

Let the stochastic process $\tilde{X}(t)$ denoting the state of the new system at any time instant t, and one has:

$$\tilde{X}(t)=\begin{cases}1, & \text{the new system is in the operating state at time } t;\\ 0, & \text{the new system is in the failed state at time } t.\end{cases}$$

It can be figured out that the Markov property is not held in the new system. In fact, given the present state of the new system, its future is related to its past and is not independent. For instance, take the repair time point of $\tau_1(\tau_1<\tau)$ as an example, the system state is failed for the original model and is operating for the new model. The future state for the new system is related to the repair time it has been processed. Thus, it is a stochastic process without Markov property or memorylessness.

3. Possibilistic availability analysis

In this section, the instantaneous possibilistic availability for the new system is discussed. Let $A_\Pi(t)$ and $\tilde{A}_\Pi(t)$ denote the instantaneous possibilistic availability of the original system and the new system, respectively.

$$\begin{aligned}A_\Pi(t)&=\Pi(\text{the original system is operating at time } t)\\&=\Pi\{X(t)=1\}\end{aligned} \tag{1}$$

$$\begin{aligned}\tilde{A}_\Pi(t)&=\Pi(\text{the new system is operating at time } t)\\&=\Pi\{\tilde{X}(t)=1\}\end{aligned} \tag{2}$$

In the first subsection, instantaneous possibilistic availability of the original system $A_\Pi(t)$ is mathematically derived. Relationship between $A_\Pi(t)$ and $\tilde{A}_\Pi(t)$ are stated in the second subsection, and the expression of $\tilde{A}_\Pi(t)$ as well.

3.1. Mathematical derivation for $A_\Pi(t)$

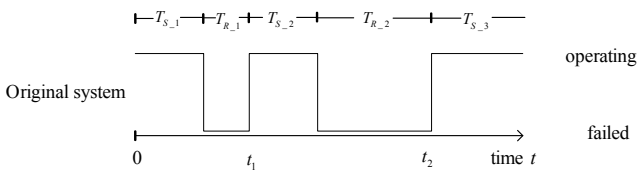


Fig.2. Process progress for the original system

As stated in Eq. (1), it is defined as the possibility that the original system is operating at time *t*. Taking account into the system process progress depicted in Fig.2, $t_1, t_2, \cdots$ are regenerative points, since the failed unit can be regarded "as good as new" after repair. Suppose that $Z_i = T_{S\_i} + T_{R\_i}$, within which T_{S_i} and T_{R_i} denote the system lifetime and repair time within i th period respectively. Thus, $\{Z_i,\ i=1,2,\cdots\}$ is a sequence of variable with independent identical distribution. Towards the event of system operating at time t ($X(t)=1$), a restatement of the event can be addressed as follows:

$$\begin{aligned}\{X(t)=1\}=&\{T_{S\_1}>t\}\cup\{Z_1<t<Z_1+T_{S\_2}\}\cup\{Z_1+Z_2<t<Z_1+Z_2+T_{S\_3}\}\cup\cdots\\&\cup\left\{\sum_{j=1}^{i}Z_j<t<\sum_{j=1}^{i}Z_j+T_{S\_i+1}\right\}\cup\cdots\end{aligned} \tag{3}$$

Thus, the instantaneous possibilistic availability of the original system can be represented as the possibility of formulation on the right hand side of Eq. (3). As for the right hand side of Eq. (3), it can be rewritten as:

$$\{T_{S\_1}>t\}\cup\bigcup_{i=1}^{\infty}\left\{\sum_{j=1}^{i}Z_j<t<\sum_{j=1}^{i}Z_j+T_{S\_i+1}\right\}$$

Additionally, based on the memoryless property of each point for $\{Z_i,\ i=1,2,\cdots\}$ and Dubois and Prade's idea [6] in defining the conditional possibility of events ($\Pi(A\cap B)=\Pi(A|B)*\Pi(B)$), it can be inferred that, for $*=\min$,

$$\begin{aligned}\Pi(Z_1<t<Z_1+T_{S\_2})&=\Pi(T_{S\_2}>t-Z_1, 0<Z_1<t)\\&=\min\{\Pi(T_{S\_2}>t-Z_1|0<Z_1<t),\Pi(0<Z_1<t)\}\\&=\min\left\{\sup_{u\in(0,t)}\Pi(T_{S\_2}>t-Z_1|Z_1=u),\sup_{u\in(0,t)}\Pi(Z_1=u)\right\}\\&=\min\left\{\sup_{u\in(0,t)}\Pi(T_{S\_2}>t-u),\sup_{u\in(0,t)}\Pi(Z_1=u)\right\}\end{aligned} \tag{4}$$

Similarly,

$$\Pi\left(Z_1+Z_2<t<Z_1+Z_2+T_{S\_3}\right)=\Pi\left(T_{S\_3}>t-Z_1-Z_2, 0<Z_1+Z_2<t\right) \tag{5}$$
$$=\min\left\{\Pi\left(T_{S\_3}>t-Z_1-Z_2 \mid 0<Z_1+Z_2<t\right), \Pi\left(0<Z_1+Z_2<t\right)\right\}$$
$$=\min\left\{\sup_{u\in(Z_1,t)}\Pi\left(T_{S\_3}>t-u\right), \sup_{u\in(Z_1,t)}\Pi\left(Z_1+Z_2=u\right)\right\}$$

$$\Pi\left(\sum_{j=1}^{i}Z_j<t<\sum_{j=1}^{i}Z_j+T_{S\_i+1}\right)=\Pi\left(T_{S\_i+1}>t-\sum_{j=1}^{i}Z_j, \sum_{j=1}^{i-1}Z_j<\sum_{j=1}^{i}Z_j<t\right) \tag{6}$$
$$=\min\left\{\sup_{u\in\left(\sum_{j=1}^{i-1}Z_j,t\right)}\Pi\left(T_{S\_i+1}>t-u\right), \sup_{u\in\left(\sum_{j=1}^{i-1}Z_j,t\right)}\Pi\left(\sum_{j=1}^{i}Z_j=u\right)\right\}$$

It can be deduced from possibility theory and the last formulation of Eqs. (4-6) that the sequence

$$\Pi\left(Z_1<t<Z_1+T_{S\_2}\right), \Pi\left(Z_1+Z_2<t<Z_1+Z_2+T_{S\_3}\right), \cdots, \Pi\left(\sum_{j=1}^{i}Z_j<t<\sum_{j=1}^{i}Z_j+T_{S\_i+1}\right),$$

is decreasing. For instance, as for the items involved within the last formulation of Eq. (5), boundary for variable u is narrower than that in Eq. (4).

Thus, we have,

$$A_\Pi(t)=\Pi\left\{X(t)=1\right\} \tag{7}$$
$$=\Pi\left\{\left\{T_{S\_1}>t\right\}\cup\bigcup_{i=1}^{\infty}\left(\sum_{j=1}^{i}Z_j<t<\sum_{j=1}^{i}Z_j+T_{S\_i+1}\right)\right\}$$
$$=\Pi\left(T_{S\_1}>t\right)\vee\Pi\left\{\bigcup_{i=1}^{\infty}\left(\sum_{j=1}^{i}Z_j<t<\sum_{j=1}^{i}Z_j+T_{S\_i+1}\right)\right\}$$
$$=\Pi\left(T_{S\_1}>t\right)\vee\Pi\left(Z_1<t<Z_1+T_{S\_2}\right)$$
$$=\Pi\left(T_{S\_1}>t\right)\vee\min\left\{\sup_{u\in(0,t)}\left[\Pi\left(T_{S\_2}>t-u\right), \Pi\left(Z_1=u\right)\right]\right\}$$

Therefore, given the possibility distribution of the variables, instantaneous possibilistic availability for the original system can be derived.

Likewise, the event of system under failure state at time t can be represented by specifying that $Z_0=0$.

$$\left\{X(t)=0\right\}=\left\{T_{S\_1}\le t\le Z_1\right\}\cup\left\{Z_1+T_{S\_2}\le t\le Z_1+Z_2\right\}\cup\cdots\cup\left\{\sum_{j=1}^{i}Z_j+T_{S\_i+1}\le t\le\sum_{j=1}^{i+1}Z_j\right\}\cup\cdots$$
$$=\bigcup_{i=0}^{\infty}\left[\sum_{j=0}^{i}\left(Z_j+T_{S\_j+1}\right)\le t\le\sum_{j=0}^{i}Z_{j+1}\right] \tag{8}$$

Respectively, possibility of the first three items within the first formulation of Eq. (8) is restated as follows:

$$\Pi\left(T_{S\_1}\le t\le Z_1\right)=\Pi\left(T_{S\_1}\le t, Z_1\ge t\right) \tag{9}$$

$$\Pi\left(Z_1+T_{S\_2}\le t\le Z_1+Z_2\right)=\Pi\left(T_{S\_2}\le t-Z_1, Z_2\ge t-Z_1\right) \tag{10}$$

$$\Pi\left(\sum_{j=1}^{i}Z_j+T_{S\_i+1}\le t\le\sum_{j=1}^{i+1}Z_j\right)=\Pi\left(T_{S\_i+1}\le t-\sum_{j=1}^{i}Z_j, Z_{i+1}\ge t-\sum_{j=1}^{i}Z_j\right) \tag{11}$$

Since the sequence of variables $\left\{T_{S\_i},\ i=1,2,\cdots\right\}$ and $\left\{Z_i,\ i=1,2,\cdots\right\}$ are with independent identical distribution respectively, we can deduce the decreasing trend for the sequence of $\Pi\left(T_{S\_1}\le t\le Z_1\right)$, $\Pi\left(Z_1+T_{S\_2}\le t\le Z_1+Z_2\right), \cdots$ from Eqs. (9-11). Thus, similarly to that inferred in Eq. (7), we have:

$$\Pi\left\{X(t)=0\right\}$$
$$=\Pi\left\{\bigcup_{i=0}^{\infty}\left(\sum_{j=0}^{i}\left(Z_j+T_{S\_j+1}\right)\le t\le\sum_{j=0}^{i}Z_{j+1}\right)\right\}$$
$$=\Pi\left\{\bigcup_{i=0}^{\infty}\left(\sum_{j=1}^{i}Z_j\le t\le\sum_{j=1}^{i}Z_j+T_{S\_i+1}\right)\right\} \tag{12}$$
$$=\Pi\left(T_{S\_1}\le t\le Z_1\right)$$
$$=\Pi\left(T_{S\_1}\le t, Z_1\ge t\right)$$

Actually, there is a property for items $\Pi\left\{X(t)=0\right\}$ and $\Pi\left\{X(t)=1\right\}$. Since the union of event $\left\{X(t)=0\right\}$ and event $\left\{X(t)=1\right\}$ comes to the universe, it can be inferred from the possibility theory that $\Pi\left\{X(t)=0\right\}\vee\Pi\left\{X(t)=1\right\}=1$. Furthermore given a critical time t, either $\Pi\left\{X(t)=0\right\}$ or $\Pi\left\{X(t)=1\right\}$ would be the value of 1. Simply speaking, if the system possibilistic availability at time t is not with the value of 1, then the possibility of system being failed at time t would be 1. On the contrary, if the system is not with full possibility to be failed at time t, the system possibilistic availability is 1 at time t.

3.2. Mathematical derivation for $\tilde{A}_\Pi(t)$

Based on the assumption that the new system is operating if the original system is operating, instantaneous possibilistic availability for the new system can be expressed as

$$\tilde{A}_\Pi(t)=\Pi\left(\text{the new system is operating at time } t\right)$$
$$=\Pi\left\{\tilde{X}(t)=1\right\}$$
$$=\Pi\left\{\left[\tilde{X}(t)=1, X(t)=0\right]\vee\left[\tilde{X}(t)=1, X(t)=1\right]\right\}$$
$$=\Pi\left\{\left[\tilde{X}(t)=1, X(t)=0\right]\vee\left[X(t)=1\right]\right\} \tag{13}$$
$$=\Pi\left(\tilde{X}(t)=1, X(t)=0\right)\vee\Pi\left(X(t)=1\right)$$
$$=\Pi\left(\tilde{X}(t)=1, X(t)=0\right)\vee A_\Pi(t)$$

Hence, in order to figure out the instantaneous possibilistic availability for the new system, we only need to analyze $\Pi\left(\tilde{X}(t)=1, X(t)=0\right)$ which representing the possibility that the new system is operating while original system is failed at time t. On the basis of the model assumptions, two cases under the circumstance of new system operating while original system failed are obtained: 1) the original system is under repair at time t, and this repair time is no

longer than τ ; and 2) the original system is under repair at time *t* and repair time is longer than τ . At the meantime, the repair have lasted no longer than τ until time *t*.

In the sequel, two scenarios in terms of threshold value are discussed: 1) τ is a constant, and 2) τ is a non-negative random variable.

A. Constant critical repair time

Assuming that the threshold τ is given as a nonnegative constant, it would be distinct to compare the magnitude of *t* and τ . Therefore, the formulation of $\Pi\left(\tilde{X}(t)=1, X(t)=0\right)$ can be represented as follows:

If $t \le \tau$, then

$$\Pi\left(\tilde{X}(t)=1, X(t)=0\right)=\Pi\left(X(t)=0\right)$$

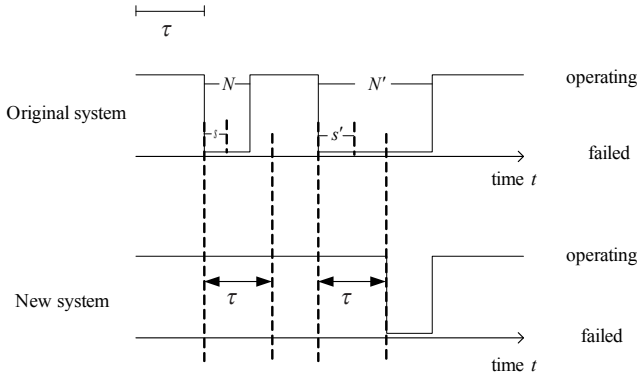


Fig.3. The new system is operating while original system is failed

If $t > \tau$, considering the two cases under the circumstance of new system operating while original system failed which are depicted in Fig.3, then

$$\begin{aligned}
&\Pi\left(\tilde{X}(t)=1, X(t)=0\right) \\
&= \min\left\{1, \int_0^{\tau} A_{\Pi}(t-s)\cdot w_{10}(t)\cdot\left[\Pi\left(s<T_R<\tau\right)\vee\Pi\left(s<T_R, T_R\ge\tau\right)\right]ds\right\} \\
&= \min\left\{1, \int_0^{\tau} A_{\Pi}(t-s)\cdot w_{10}(t)\cdot\Pi\left(s<T_R\right)ds\right\}
\end{aligned}$$

in which $w_{10}(t)$ denotes the membership function of transiting from State 1 to State 0 at time *t*. Therefore, system instantaneous possibilistic availability can be expressed as:

$$\begin{aligned}
\tilde{A}_{\Pi}(t) &= \Pi\left(\tilde{X}(t)=1, X(t)=0\right)\vee A_{\Pi}(t) \\
&= \begin{cases} \Pi\left(X(t)=0\right)\vee A_{\Pi}(t) & t\le\tau \\ A_{\Pi}(t)\vee\int_0^{\tau} A_{\Pi}(t-s)\cdot w_{10}(t)\cdot\Pi\left(s<N\right)ds & t>\tau \end{cases}
\end{aligned} \tag{14}$$

B. Random critical repair time

Suppose that the threshold is given as a random variable following distribution function of $H(\tau)$, then

$$\begin{aligned}
&\Pi\left(\tilde{X}(t)=1, X(t)=0\right) \\
&= \int_0^{\infty}\int_0^{\min(t,\tau)} A(t-s)w_{10}\cdot\left[\Pi\left(s<T_R<\tau\right)\vee\Pi\left(s<T_R, T_R\ge\tau\right)\right]dsdH(\tau) \\
&= \int_t^{\infty}\Pi\left(X(t)=0\right)dH(\tau)+\int_0^{t}\int_0^{\tau} A_{\Pi}(t-s)\cdot w_{10}(t)\cdot\Pi\left(s<T_R\right)dsdH(\tau) \\
&= \Pi\left(X(t)=0\right)\bar{H}_{\tau}(t)+\int_0^{t}\int_0^{\tau} A_{\Pi}(t-s)\cdot w_{10}(t)\cdot\Pi\left(s<T_R\right)dsdH(\tau)
\end{aligned} \tag{15}$$

The last equation holds because of $\int_t^{\infty} dH(\tau)=P(\tau\ge t)=1-H_{\tau}(t)=\bar{H}_{\tau}(t)$. Similarly, system instantaneous possibilistic reliability can be expressed making use of $\tilde{A}_{\Pi}(t)=\Pi\left(\tilde{X}(t)=1, X(t)=0\right)\vee A_{\Pi}(t)$. In which, $A_{\Pi}(t)$ can be figured out taking advantage of Eq. (7).

4. Numerical example

In this section, a numerical example is shown to compare the instantaneous possbilistic reliability between the original system and the new system. Based on the assumption that the repaired component can be restored into "as good as new", each sequence of variables $\{T_{S\_i},\ i=1,2,\cdots\}$ and $\{T_{R\_i},\ i=1,2,\cdots\}$ are with independent identical distribution respectively. Generally, we use variable T_S as a denotation for each variable in the sequence of $\{T_{S\_i},\ i=1,2,\cdots\}$, and simplify each variable within $\{T_{R\_i},\ i=1,2,\cdots\}$ as T_R , respectively. Similarly, the sequence of $\{Z_i,\ i=1,2,\cdots\}$ is simplified as Z in which $Z=T_S+T_R$.

Suppose that the system lifetime T_S follows deflection minor type possibility distribution $\pi_{T_S}(u)$ in which $t_1=100(day)$, and the repair time T_R follows deflection minor type possibility distribution $\pi_{T_R}(u)$ in which $t_2=1(day)$:

$$\pi_{T_S}(t_S)=\begin{cases} 1 & t_S\le t_1 \\ \exp\left\{-\frac{1}{2}\left(\frac{t_S-t_1}{t_1}\right)\right\} & t_S>t_1 \end{cases} \tag{16}$$

$$\pi_{T_R}(t_R)=\begin{cases} 1 & t_R\le t_2 \\ \exp\left\{-\frac{1}{2}\left(\frac{t_R-t_2}{t_2}\right)\right\} & t_R>t_2 \end{cases} \tag{17}$$

Now we come to the possibility distribution for Z . According to the possibility theory and the independence between T_S and T_R ($\pi_{(T_S,T_R)}(t_S,t_R)=\min\left(\pi_{T_S}(t_S),\pi_{T_R}(t_R)\right)$), we have:

$$\begin{aligned}
\pi_Z(z)=\pi_{T_S+T_R}(z) &= \bigvee_{t_S+t_R=z}\pi_{(T_S,T_R)}(t_S,t_R) \\
&= \bigvee_{t_S+t_R=z}\left[\min\left(\pi_{T_S}(t_S),\pi_{T_R}(t_R)\right)\right]
\end{aligned} \tag{18}$$

For the expression of instantaneous possibilistic reliability for original system $A_{\Pi}(t)$, refer to details in appendix. Here in this exam-

ple, τ is given the value of 2(*day*) and $w_{10}(t)$ is assumed to be a sub-function defined in Eq. (19) . As a matter of fact, according to Eq. (29) the system is with full possibility to be operating when $t \in [0, 2t_1 + t_2]$, which means it is not likely for the occurrence of system state change from operating to failed. Conversely, it is more likely for the system to become failed from operating state when $t \in (2t_1 + t_2, \infty)$. It is because the system possibilistic availability is decreasing.

$$w_{10}(t) = \begin{cases} 0.15 & t \in [0, 2t_1 + t_2] \\ 0.85 & t \in (2t_1 + t_2, \infty) \\ 1 & t = \infty \end{cases} \tag{19}$$

According to the analysis in the previous section, instantaneous possibilistic reliability for new system can be figured out taking advantage of Eq. (14).

$$\begin{aligned}
\tilde{A}_{\Pi}(t) &= \begin{cases} 1 & t \le 2(day) \\ A_{\Pi}(t) \vee \min\left\{1, \int_0^2 A_{\Pi}(t-s) \cdot w_{10}(t) \cdot \Pi(s < N) ds\right\} & t > 2(day) \end{cases} \\
&= \begin{cases} 1 & t \le 201(d) \\ \exp\left\{-\frac{1}{2}\left(\frac{t-201}{200}\right)\right\} \vee \min\left\{1, 0.85\int_0^2 \exp\left\{-\frac{1}{2}\left(\frac{t-201-s}{200}\right)\right\} \cdot \exp\left\{-\frac{1}{2}(s-1)\right\} ds\right\} & t > 201(d) \end{cases} \\
&= \begin{cases} 1 & t \le 201(d) \\ \exp\left\{-\frac{1}{2}\left(\frac{t-201}{200}\right)\right\} \vee \min\left\{1, 0.85 \cdot \exp\left\{-\frac{t-401}{400}\right\} \cdot \int_0^2 \exp\left\{-\frac{199}{400} s\right\} ds\right\} & t > 201(d) \end{cases} \\
&= \begin{cases} 1 & t \le 201(d) \\ \exp\left\{-\frac{1}{2}\left(\frac{t-201}{200}\right)\right\} \vee \min\left\{1, 1.0139 \cdot \exp\left\{-\frac{t-401}{400}\right\}\right\} & t > 201(d) \end{cases}
\end{aligned}$$

From Fig. 4, it can be figured out that the instantaneous possibilistic availability for the new system is higher than that for the original system. It should be this situation due to the emergence of neglected or omitted failure in the new system. At the meantime, it is shown from Fig.4, $\tilde{A}_{\Pi}(t) = 1$ holds for the range of $t \le 2t_1 + t_2$. It seems surprisingly for such a consequence. In fact, the result indicates that for the range of $t \le 2t_1 + t_2$, system instantaneous availability is capable to be 1.

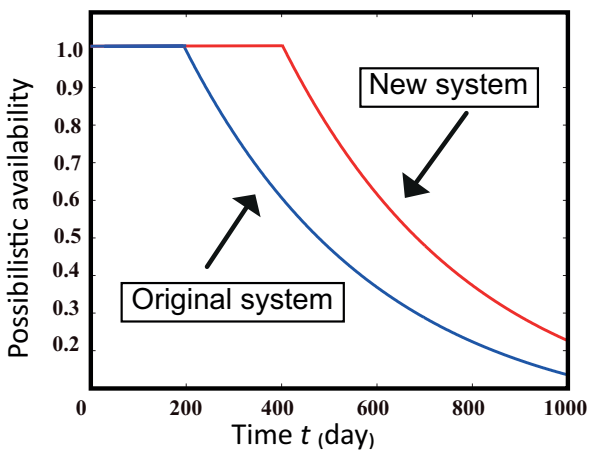


Fig. 4. The curves of the possibilitic availabilities for the new system and the original system

5. Conclusion

On the basis of some practical problems in system maintenances, a new single-unit repairable system is proposed in this paper. In such a new system, a short repair may lead to a system failure. Given a critical value, if the repair time is less than the value, the repair interval can be omitted, i.e., the failure effect is omitted. If the repair time is longer than the value, then the system is considered to remain in the operating state from the initial stage of the repair till the end of the repair threshold, i.e., the failure effect is delayed.

Considering the epistemic uncertainty which widely exists in practical engineering, system possibilistic availability is analyzed based on possibility theory. As for the difference between probability theory and possibility theory, one may be stunned by the result that the possibilistic availability is with a high value. Possibility denotes the capability for the system. Thus, it is with a higher value compared with probability.

We consider the very simple system 'single-unit system' in the paper. More complicated system will be discussed for the application in practical engineering in the future. Furthermore, various indices will be considered to offer more information for the system.

Appendix

Calculation for instantaneous possibilistic reliability of original system $A_{\Pi}(t)$

In order to figure out the instantaneous possibilistic reliability of the original system which is presented in Eq. (7), four phases for time *t* are distinguished as follows:

I If $t \in [0, t_1]$, it can be easily deduced from the last equation in Eq. (7) that $\Pi(T_S > t) = \Pi(T_S = t_1) = 1$, thus, $A_{\Pi}(t) = 1$.

II If $t \in (t_1, t_1 + t_2]$, thus it can be obtained from $u \in (0, t)$ that $u \in (0, t_1 + t_2)$. Since that $t_S < t_1$ results in $\pi_{T_S}(t_S) = 1$ and $t_R < t_2$ leads to the consequence of $\pi_{T_R}(t_R) = 1$, we have:

$$\begin{aligned}
A_{\Pi}(t) &= \Pi(T_S > t) \vee \min\left\{\sup_{u \in (0,t)} \left[\Pi(T_S > t-u), \Pi(Z = u)\right]\right\} \\
&= \Pi(T_S > t) \vee \min\left\{\sup_{\substack{u \in (0,t) \\ t \in (t_1, t_1+t_2]}} \left[\Pi(T_S > t-u), \bigvee_{t_S + t_R = u} \left[\min\left(\pi_{T_S}(t_S), \pi_{T_R}(t_R)\right)\right]\right]\right\} \\
&= \Pi(T_S > t) \vee \min\left\{\sup_{\substack{u \in (0,t) \\ t \in (t_1, t_1+t_2]}} \left[\Pi(T_S > t-u), \bigvee_{t_S < t_1, t_R < t_2} \left[\min\left(\pi_{T_S}(t_S), \pi_{T_R}(t_R)\right)\right]\right]\right\} \\
&= \Pi(T_S > t) \vee \sup_{\substack{u \in (0,t) \\ t \in (t_1, t_1+t_2]}} \left[\Pi(T_S > t-u)\right] \\
&= \sup_{\substack{u \in (0,t) \\ t \in (t_1, t_1+t_2]}} \left[\Pi(T_S > t-u)\right] \\
&= 1
\end{aligned} \tag{20}$$

In fact, as long as there is an opportunity for *u* and *t* satisfy that $t - u \le t_1$, the possibilistic availability at time t ($t \in (t_1, t_1 + t_2]$) is 1.

III If $t \in (t_1+t_2, 2t_1+t_2]$,

1) When $u \in (0, t_1+t_2)$, it can be similarly inferred as Eq. (19)

$$\begin{aligned} A_\Pi(t) &= \Pi(T_s > t) \vee \min\left\{ \sup_{u\in(0,t_1+t_2)} \left[\Pi(T_s > t-u), \bigvee_{t_S<t_1, t_R<t_2} \left[\min\left(\pi_{T_s}(t_S), \pi_{T_R}(t_R)\right)\right]\right]\right\} \\ &= \Pi(T_s > t) \vee \sup_{u\in(0,t_1+t_2)} \left[\Pi(T_s > t-u)\right] \\ &= \Pi(T_s > t) \vee \Pi(T_s > t-t_1-t_2) \\ &= \Pi(T_s > t-t_1-t_2) \\ &= 1 \end{aligned} \tag{21}$$

The last equal mark holds for that the value range of time t is limited by the upper bound of $2t_1+t_2$. Or else, $\Pi(T_S > t-t_1-t_2) = 1$ doesn't hold.

2) When $u \in [t_1+t_2, t)$,

a) If $t_S \le t_1, t_R \ge t_2$, then $\pi_{T_S}(t_S) = 1$ and $\pi_{T_R}(t_R) = 1$. Apparently, it holds for $\min\left(\pi_{T_S}(t_S), \pi_{T_R}(t_R)\right) = \pi_{T_R}(t_R)$.

$$\begin{aligned} A_\Pi(t) &= \Pi(T_S > t) \vee \min\left\{ \sup_{u\in[t_1+t_2,t)} \left[\Pi(T_S > t-u), \Pi(Z=u)\right]\right\} \\ &= \Pi(T_S > t) \vee \min\left\{ \sup_{u\in[t_1+t_2,t)} \left[\Pi(T_S > t-u), \bigvee_{t_S\le t_1, t_R\ge t_2} \left[\min\left(\pi_{T_S}(t_S), \pi_{T_R}(t_R)\right)\right]\right]\right\} \\ &= \Pi(T_S > t) \vee \min\left\{ \sup_{u\in[t_1+t_2,t)} \left[\Pi(T_S > t-u), \left(\pi_{T_R}(u-t_1)\right)\right]\right\} \\ &= \Pi(T_S > t) \vee \min_{u=t_1+t_2} \left[\Pi(T_S > t-u), \left(\pi_{T_R}(u-t_1)\right)\right] \\ &= \Pi(T_S > t) \vee 1 \\ &= 1 \end{aligned} \tag{22}$$

b) If $t_R \le t_2, t_S \ge t_1$, it can be similarly derived as in Eq. (21).

$$\begin{aligned} A_\Pi(t) &= \Pi(T_S > t) \vee \min\left\{ \sup_{u\in[t_1+t_2,t)} \left[\Pi(T_S > t-u), \bigvee_{t_S\ge t_1, t_R\le t_2} \left[\min\left(\pi_{T_S}(t_S), \pi_{T_R}(t_R)\right)\right]\right]\right\} \\ &= \Pi(T_S > t) \vee \min_{u=t_1+t_2} \left[\Pi(T_S > t-u), \left(\pi_{T_S}(u-t_2)\right)\right] \\ &= \Pi(T_S > t) \vee 1 \\ &= 1 \end{aligned} \tag{23}$$

IV If $t \in (2t_1+t_2, \infty)$,

1) When $u \in (0, t_1+t_2)$,

$$\begin{aligned} A_\Pi(t) &= \Pi(T_S > t) \vee \min\left\{ \sup_{u\in(0,t_1+t_2)} \left[\Pi(T_S > t-u), \bigvee_{t_S<t_1, t_R<t_2} \left[\min\left(\pi_{T_S}(t_S), \pi_{T_R}(t_R)\right)\right]\right]\right\} \\ &= \Pi(T_S > t) \vee \sup_{u\in(0,t_1+t_2)} \left[\Pi(T_S > t-u)\right] \\ &= \Pi(T_S > t) \vee \Pi(T_S > t-t_1-t_2) \\ &= \Pi(T_S > t-t_1-t_2) \end{aligned} \tag{24}$$

2) When $u \in [t_1+t_2, t)$,

a) If $t_S \le t_1, t_R \ge t_2$,

$$\begin{aligned} A_\Pi(t) &= \Pi(T_S > t) \vee \min\left\{ \sup_{u\in[t_1+t_2,t)} \left[\Pi(T_S > t-u), \left(\pi_{T_R}(u-t_1)\right)\right]\right\} \\ &= \Pi(T_S > t) \vee \min_{u=t-t_1} \left[\Pi(T_S > t-u), \left(\pi_{T_R}(u-t_1)\right)\right] \\ &= \Pi(T_S > t) \vee \Pi(T_R > t-2t_1) \end{aligned} \tag{25}$$

b) If $t_R \le t_2, t_S \ge t_1$,

$$\begin{aligned} A_\Pi(t) &= \Pi(T_S > t) \vee \min\left\{ \sup_{u\in[t_1+t_2,t)} \left[\Pi(T_S > t-u), \left(\pi_{T_S}(u-t_2)\right)\right]\right\} \\ &= \Pi(T_S > t) \vee \min_{u=\frac{t+t_2}{2}} \left[\Pi(T_S > t-u), \left(\pi_{T_S}(u-t_2)\right)\right] \\ &= \Pi(T_S > t) \vee \Pi\left(T_S > \frac{t-t_2}{2}\right) \\ &= \Pi\left(T_S > \frac{t-t_2}{2}\right) \end{aligned} \tag{26}$$

Thus, if $t \in [0, 2t_1+t_2]$, then $A_\Pi(t) = 1$. If $t \in (2t_1+t_2, \infty)$, then

$$\begin{aligned} A_\Pi(t) &= \Pi(T_S > t-t_1-t_2) \vee \Pi(T_S > t) \vee \Pi(T_R > t-2t_1) \vee \Pi\left(T_S > \frac{t-t_2}{2}\right) \\ &= \Pi(T_R > t-2t_1) \vee \Pi\left(T_S > \frac{t-t_2}{2}\right) \end{aligned} \tag{27}$$

Additionally together with Eqs. (16, 17), it is able to go a step further for the simplification of the expression ahead. Moreover, the last equation within the following formulation is deduced for that $t_1 > t_2$.

$$\begin{aligned} A_\Pi(t) &= \Pi(T_R > t-2t_1) \vee \Pi\left(T_S > \frac{t-t_2}{2}\right) \\ &= \exp\left\{-\frac{1}{2}\left(\frac{t-2t_1-t_2}{t_2}\right)\right\} \vee \exp\left\{-\frac{1}{2}\left(\frac{t-2t_1-t_2}{2t_1}\right)\right\} \\ &= \exp\left\{-\frac{1}{2}\left(\frac{t-2t_1-t_2}{2t_1}\right)\right\} \end{aligned} \tag{28}$$

Generally, instantaneous possibilistic reliability of original system can be expressed as

$$A_\Pi(t) = \begin{cases} 1 & t \in [0, 2t_1+t_2] \\ \exp\left\{-\frac{1}{2}\left(\frac{t-2t_1-t_2}{2t_1}\right)\right\} & t \in (2t_1+t_2, \infty) \end{cases} \tag{29}$$

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China under the contract number 51075061 and the Specialized Research Fund for the Doctoral Program of Higher Education of China under the contract number 20090185110019.

References

1. Ascher H E. Repairable Systems Reliability. In: Encyclopedia of Statistics in Quality and Reliability, New York: John Wiley & Sons, 2008.
2. Bao X Z, Cui L R. An analysis of availability for series Markov repairable system with neglected or delayed failures. IEEE Transactions on Reliability 2010; 59(4):734-743.
3. Baraldi P, Zio E. A combined Monte Carlo and possibilistic approach to uncertainty propagation in event tree analysis. Risk Analysis 2008; 28:1309-1325.

4. Cao J H, Cheng K. Introduction to Reliability Mathematics. Science Press, Beijing, 1986.
5. Cao J H, Wu Y H. Reliability analysis of a multi state repairable system with a replaceable repair facility. Acta Mathematicae Applicate Sinica 1988; 4(2): 113-121.
6. Dubois D, Prade H. Possibility Theory. Plenum Press, New York.1988.
7. He L P, Qu F Z. Possibilistic entropy-based measure of importance in fault tree analysis. Journal of Systems Engineering and Electronics 2009; 20: 434-444.
8. Hunter A, Liu W R. A context-dependent algorithm for merging uncertain information in possibility theory. IEEE Transactions on Systems Man and Cybernetics: Part A-Systems and Humans, 2008; 38:1385-1397.
9. Jia Z Y, Kang R. Wang L C, Wang N C. Availability research on K-out-of-N: G systems with repair time omission. Advanced Materials Research, Materials and Product Technologies II, 2010; 118-120: 342-347.
10. Li H X, Meng X Y, Li N. An availability analysis of series repairable system with repair time omission. Journal of Yanshan University 2007; 31(6): 542–544.
11. Liu Y, Tang W S, Zhao R Q. Reliability and mean time to failure of unrepairable systems with fuzzy random lifetimes. IEEE Transactions on Fuzzy Systems 2007; 15(5): 1009-1026.
12. Krivtsov V V. Practical extensions to NHPP application in repairable system reliability analysis. Reliability Engineering & System Safety 2007; 92(5): 560-562.
13. Mourelatos Z P, Zhou J. Reliability estimation and design with insufficient data based on possibility theory. AIAA Journal 2005; 43:1696-1705.
14. Oberkampf W L, Helton J C, Joslyn C A, Wojtkiewicz S F, Ferson S. Challenge problems: Uncertainty in system response given uncertain parameters. Reliability Engineering and System Safety 2004; 85:11-19.
15. Pang Y, Huang H Z, Liu Y, Miao Q, Wang Z L. Reliability analysis of a repairable parallel system with repair time omission. Proceedings of 8th International Conference on Reliability, Maintainability and Safety (2009 ICRMS), 2009: 41-45.
16. Podofillini L, Zio E, Mercurio D, Dang V N. Dynamic safety assessment: Scenario identification via a possibilistic clustering approach. Reliability Engineering and System Safety 2010; 95: 534-549.
17. Yuan L, Xu J. An optimal replacement policy for a repairable system based on its repairman having vacations. Reliability Engineering & System Safety 2011; 96(7): 868-875.
18. Zhang Y L, Wang G J. A deteriorating cold standby repairable system with priority in use. European Journal of Operational Research 2007; 183(1): 278-295.
19. Zhang Y L, Wu S M. Reliability analysis for a k/n (F) system with repairable repair-equipment. Applied Mathematical Modelling 2009; 33(7):3052-3067.
20. Zheng Z H, Cui L R, Hawkes A G. A study on a single-unit Markov repairable system with repair time omission. IEEE Transaction on Reliability 2006; 55(2): 182–188.
21. Zheng Z H, Cui L R. Availability analysis of parallel repairable system with omitted failures. Journal of Beijing Institute of Technology (English Edition) 2009; 18(1): 117-120.
22. Zheng Z H. A Study on a single-unit Markov repairable system with effects neglected or delayed of failures. Master thesis, Beijing Institute of Technology, 2006.

Yu PANG, Ph.D. candidate
Prof. Hong-Zhong HUANG, Ph.D.
Ning-Cong XIAO, Ph.D. candidate
Associate Prof. Yu LIU, Ph.D.
Yanfeng LI, Ph.D. candidate
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:
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.

Robert PILCH
Jan SZYBKA
Zdzisław BRONIEC

DETERMINING OF HOT WATER-PIPE EXPLOITATION TIME ON THE BASIS OF LIMITING STATES

WYZNACZANIE CZASU EKSPLOATACJI CIEPŁOCIĄGU NA PODSTAWIE IDENTYFIKACJI STANÓW GRANICZNYCH*

Methodology for determining the limiting time of the hot water-pipeline exploitation has been described in the present study. Pitting corrosion causing local reduction of the hot water-pipeline wall thickness was assumed as the basis of the limiting time determining. Three limiting states influencing the hot water-pipeline strength were taken into consideration in the executed analysis. It was estimated that the wall thickness distribution is consistent with normal distribution as well as that exceeding of the hot water-pipeline wall thickness limiting values within given probability level is a basis for the exploitation time determining.

***Keywords**: hot water-pipeline, corrosion, limiting states, exploitation time.*

W artykule przedstawiono metodykę wyznaczania granicznego czasu eksploatacji rurociągu. Podstawą wyznaczenia tego czasu jest korozja wżerowa powodująca lokalne zmniejszanie grubości ścianek rurociągu. W analizie wzięto pod uwagę trzy stany graniczne decydujące o wytrzymałości rurociągu. Oszacowano, że rozkład grubości ścianek jest zgodny z rozkładem normalnym a przekroczenie granicznych wartości grubości ścianki rurociągu na zadanym poziomie prawdopodobieństwa jest podstawą do określenia czasu jego eksploatacji.

***Słowa kluczowe**: ciepłociąg, korozja, stany graniczne, czas eksploatacji.*

1. Introduction

Occurrence of the corrosion in hot water-pipelines is the main cause of the exploitation properties reduction and with time the user is forced to make costly pipe replacements. Corrosion process inside hot water-pipeline is continuous and there are practically none possibilities to prevent the process.

In the worldwide literature there are many of positions concerning estimation of the gas transfer networks, which take under consideration the corrosion phenomena [2, 5]. Problems of the corrosion of hot water-pipelines are broadly discussed in [3, 4, 12]. The procedure presented in this paper is original and the authors focused their attention to the problems of the hot water-pipeline strength, with special attention paid to pitting corrosion phenomenon. Appearance of the corrosion pits has random character and the places of their occurrence are difficult to locate.

The authors have undertaken the difficult task of determining the hot water-pipeline exploitation time with respect to reaching the limiting times caused by exceeding permissible stresses within the pipe material resulting from development of corrosion pits.

Problems described in the present study concern a specific hot water-line exploited in one of Polish mines.

2. Characteristics of the corrosion process and determining the hot water-pipeline limiting states

Hot water-pipeline which is exposed to adverse impacts of water-based chemical environment and to appearance of variable loads (pressure pulsation), including random pulsation as a subject of progressive degradation, which takes place in result of the action of chemical compounds present in water, high temperature and time.

Occurring corrosion is considered as an electro-chemical process, which takes place in continuous manner and leads to deepening of the corrosion pits on the pipe wall [14]. These changes of the depth of the corrosive losses have character of quantitative changes (they can be measured). In consequence the limiting states of the hot water-pipeline walls are exceeded, what in turn results in break-down (qualitative change) or increased probability of the break-down occurrence.

Process of the corrosion pits deepening may be treated as random process consisting of successive phases related with change of the corrosion loses height and events leading to qualitative changes [6].

Periodical measurements of the wall thickness can be considered as controlled phases of the process development. After some time of the exploitation, the assumed critical states are exceeded, what is considered as events proving partial damages leading in consequence to break-down, which is in turn considered as the process failure (hot water-pipeline).

It can be characterized as a stochastic process $N(t)$ of integer-non-negative values and continuous time, both by the distribution of number of events taking place within time intervals having length corresponding to real process duration and distribution of the lengths of intervals between occurring events.

In a case described in the present study, the electro-chemical corrosion after a certain period of the hot water-pipeline exploitation, depth of corrosion pits becomes a random variable of defined probability distribution. Taking under consideration limiting values of the hot water-pipeline wall thickness and increase of the depth of corrosion pits, we can determine percentage fraction of the corrosion pits,

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

which are dangerous for the hot water-pipeline exploitation, as well as we can assess distribution, which characterizes increasing in time probability of the failure occurrence. Assuming a particular permissible level of risk of the failure, we can determine the hot water-pipeline exploitation time.

Determining of the hot water-pipeline exploitation time is not possible without determining the hot water-pipeline strength. Resistance to failures will depend both on external loads and course of the corrosion process development.

Circumferential stresses σ are generated within light-wall pipelines exposed to action of external pressure p, which have the same value within the whole pipeline thickness h (in given cross-section) [7, 8].

$$s = \frac{r \cdot p}{h} \tag{1}$$

$$r = \frac{(D_Z + D_W)}{4}$$

where:
r – mean pipe radius,
D_z – outside pipe diameter,
D_w – inside pipe diameter.

Thin wall condition in form $\frac{h}{r} \leq 0{,}2$ is satisfied.

Strength analysis of the pipeline can be considered on the basis of required thicknesses of the pipeline wall, assuming critical strength K_r:
- yield stress $K_r = R_e$,
- immediate strength $K_r = R_m$,
- and fatigue strength limit $K_r = Z_{rj}$.

Using the formula (1) we can determine the wall thickness:

$$h = \frac{D_Z \cdot p}{2 \cdot K_r + p} \tag{2}$$

Thickness of the pipeline walls h with respect to assumed values of critical stress K_r is schematically shown in Fig. 1.

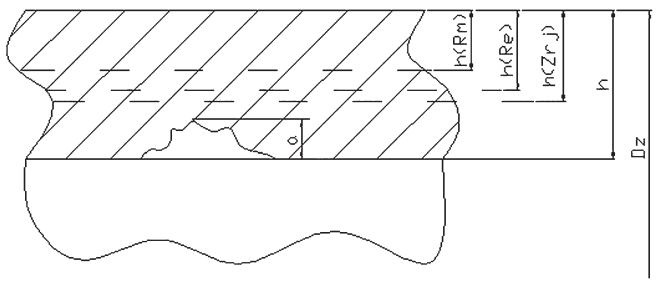


Fig. 1. Pipeline wall with example of corrosion pit and required limiting thicknesses

$h(R_m)$ – minimal permissible pipe thickness with respect to immediate strength,
$h(R_e)$ – minimal permissible pipe thickness with respect to yield stress,
$h(Z_{rj})$ – minimal permissible pipe thickness with respect to fatigue strength,
a – corrosion pit depth,
h – nominal pipe wall thickness.

3. The pipeline exploitation time calculation

One of the basic problems in the pipeline exploitation is related with assuring continuous and safe transport of the medium. In order to realize this task the user must control the pipeline condition, particularly consequent decrease of pipe strength, what can lead to the pipeline failure. Corrosion of the pipeline walls is one of factors influencing pipe weakening. If this process in continuous in time we can easily determine time, when we should undertaken suitable steps to avoid the pipeline failure.

The problem is much more complicated, if the pit corrosion occurs, as described in point 2. Change of the pipe wall thickness (and development of dip corrosion pits) is caused mostly by electro-chemical corrosion, resulting from high level of the water mineralization (presence of the great amounts of chlorides).

Beside uniform decrease of the wall thickness, some places with dip pits are observed, which can be considered as potential points of the failure occurrence – Fig. 1. Additional problem is related with the fact that the pits are usually scattered along the whole pipeline length, and they cannot be detected during the control. From the other side, control of the whole pipeline is often very costly and time-consuming, and sometimes simply impossible. Thus the user often faces the problem of determining the risk level of the pipeline exploitation, including time remaining until making decision preventing the failure, having incomplete information about the pipeline condition.

One possibility of calculation the mentioned values comprises use of probabilistic models describing wall thickness and corrosion pits distributions, including their changes in time.

Determining of the distribution of remaining wall thicknesses (or corrosion pits depth) of the tested pipeline can be executed on the basis of the executed depth measurements results. Based on the compatibility tests we can match suitable probability distribution and determine its parameters. Than taking under consideration the exploitation conditions, parameters of transported medium and their variability in time, the required limiting wall thickness value should be determined with respect to assumed critical state, as shown in 2. Having distribution of the wall thicknesses and required limiting values we can determine probability of exceeding the assumed limiting value. Wall thicknesses for normal distribution of the pipeline are shown in Fig. 2.

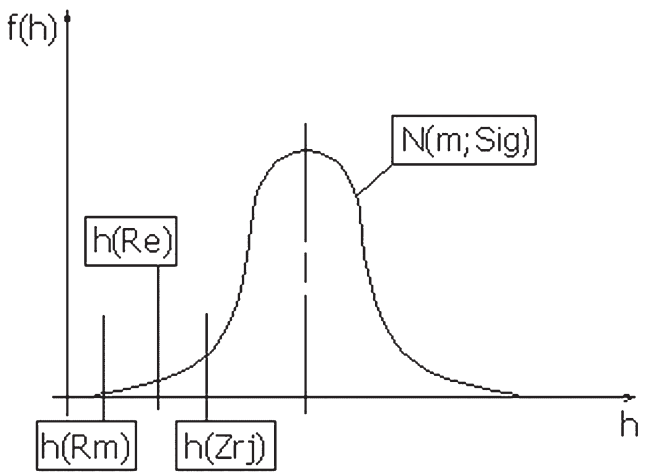


Fig. 2. Density function of the wall thicknesses distribution with required limiting thicknesses

Specified probability of exceeding the limiting value gives us information about the degree of failure risk in the moment when the wall thickness measurements were executed. However, designation of the time after which chosen critical state with assumed reliability level is reached, is also essential.

In connection with the ongoing corrosion process, determined distribution of the wall thicknesses will be gradually shifted. If you have the measurements of the wall thickness executed in various periods of the pipeline exploitation, their distributions and manner of the parameters change in time should be determined. If the executed analysis allows determining the distribution shifting manner, including character of these changes in time, we can on this basis make a prognosis of the distribution changes in time. Then using method of successive approximation we can determine time after which the wall thickness reaches limiting value on assumed probability level – Fig. 3.

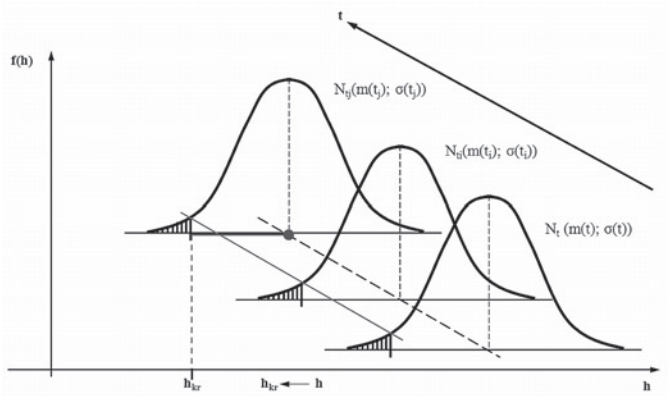


Fig. 3. Calculation of the time of limiting state on assumed probability level

4. Practical example

Weld-less R35 low carbon (P235GH) steel pipes of the diameter φ508x11 were used for the pipeline building, and basal requirements are defined in Polish standard No PN-80H-74219 [13]. The pipes were manufactured with use of the hot rolling technology, with beveled walls and calibrated endings.

Required minimal value of the pipe material yield stress in temperature of 473 [K] (200 [ºC]) should satisfy inequality $R_e \geq 185$ [MPa] where $R_m = 345$ [MPa].

Pressure of the transported medium is the basal load for the examined pipeline. Operational values of the transported medium determined on the basis of round-the-clock service oscillate within the range 1,4 ÷ 2,7 [MPa].

Exploitation of the pipeline is also threatened by variable stresses resulting from non-uniform pressure of the transported medium. Moreover, data collected from the round-the clock monitoring also prove its wavy course as well as stochastic values generating additional load impulses.

Assuming experimental relation for steel between fatigue limit Z_{rj} and immediate strength R_m ($Z_{rj} \approx 0{,}5\ R_m$), needed wall thicknesses were obtained:

$$h(R_e) = 3{,}7\ [\text{mm}],\ h(R_m) = 2{,}0\ [\text{mm}],\ h(Z_{rj}) = 3{,}9\ [\text{mm}].$$

Moreover, taking under consideration dynamic action of the pressure with factor 2 [1] in comparison to static pressure ($p_{dyn} = 2p_{stat}$) we obtain the following wall thicknesses:

$$h(R_e) = 7{,}3\ [\text{mm}],\ h(R_m) = 3{,}9\ [\text{mm}],\ h(Z_{rj}) = 7{,}8\ [\text{mm}].$$

In case of hot-water pipeline, measurements of the wall thicknesses were repeated three times during the pipeline exploitation. Wall thickness measurements were executed with use of ultrasonic thickness gauge (type DM-4DL, head DA 317). Moreover, thermovision measurements (camera type Ti25), were executed in the same points. The computer-processed results of these measurements are shown in Figure 4.

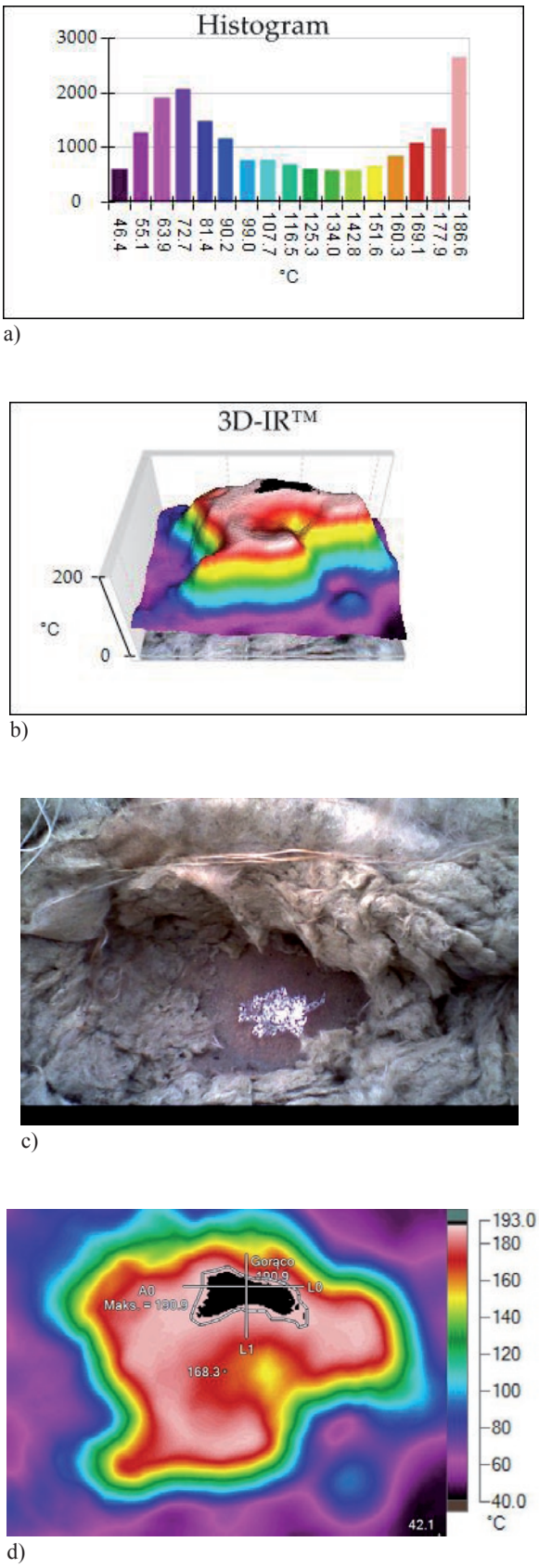


Fig. 4. Example of the thermovision measurement [11], a) image within visible range, b) image in infrared, c) histogram of given temperature value, d) three-dimensional temperature distribution on tested surface

On the basis of measurement results and calculated limiting values of the wall thickness, actual reliability level, i.e. probability of reaching limiting state resulting from possible material fatigue, as well as time after which the reliability drops to assumed limiting value amounting for 0,98, should be determined. Assumed high value of required reliability results from a significant risk of the pipeline exploitation.

The executed compatibility tests (Shapiro-Wilk test) proved that measurements of the wall thickness executed in 6, 8 and 17-th year of exploitation are characterized with normal probability distributions, which can be described consequently as: N6(12,33; 0,9), N8(11,86; 1,46), N17(10,53; 1,13). On the basis of most recent results of the wall thickness measurements executed in 17-th year of exploitation (in the year 2010), described with normal distribution N17(10,53; 1,13), probability of reaching limiting states and reliability levels corresponding to these states, have been determined. Results are shown in Table 1.

Table. 1. The pipeline reliability in the 17-th year of exploitation

	Limiting state with respect to:		
	occurrence of the material fatigue Z_{rj}	yield stress R_e	strength limit R_m
Reliability in 17-th year of exploitation	0,9921	0,9979	1

It should also be noted that mean change of the pipe wall in time has almost perfectly linear character, with factor R2=0,9933, what is shown in Fig. 5.

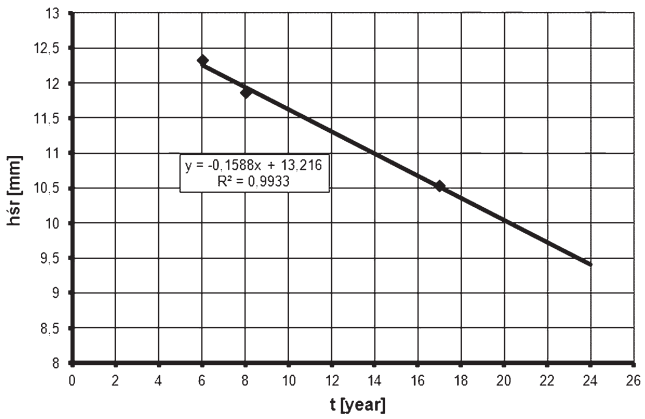


Fig. 5. Change in time of the pipe wall thickness $h_{śr}$ for tested pipeline

Determining the equation of time changes of the pipe wall thickness, which has a form:

$$h_{sr} = -0{,}1588\ t + 13{,}216\ ; \qquad (3)$$

forecasted mean values of the pipe wall thickness within slidable normal distribution for consecutive exploitation years have been calculated. Then assuming the mean invariable standard deviation in forecasted distribution, reliability levels have been determined and time after which the limiting value is reached, has been calculated (Fig. 3). The results are gathered in Table 2.

Table. 2. Reliability of the pipeline in consecutive exploitation years

Exploitation year	Forecasted $h_{śr}$	Probability of not achieving the limiting state with respect to:		
		occurrence of the material fatigue Z_{rj}	limit of yield stress R_e	limit of the strength R_m
18	10,36	0,9863	0,9958	1
18,5	10,28	0,9837	0,9949	
19	10,20	0,9807	0,9938	
19,5	10,12	0,9772	0,9925	
20	10,04	0,9733	0,9909	
20,5	9,96	0,9687	0,9891	
21	9,88	0,9635	0,9869	
21,5	9,80	0,9576	0,9844	
22	9,72	0,9510	0,9815	
22,5	9,64	0,9436	0,9782	

The obtained results prove that between 19,5 and 22,5-th exploitation year the pipe failure risk is relatively high because it exceeds value of 2% , and within this range the pipeline strength is located between limits resulting from material fatigue and plasticity. This exploitation period should be taken as a limit of safe exploitation and suitable time for the pipeline replacement. The result obtained are similar to the results of the strength tests described in works [9, 10].

Further tests aimed at better accommodation of the pipeline to real conditions indicate that the distribution of the pipeline operation until the limiting state with probability of 0,02 is reached, is compatible with the normal distribution N(28,9; 9,6). The obtained preliminary results require further verification and they will be presented in the next study.

5. Summary

Determination of the hot water-pipeline exploitation time is important for the potential user with respect to assure the production continuity. Each unexpected failure is accompanied with big economic losses and affects the production stability. Over time, the number of failures increases and the term of the pipeline replacement and the deadline for the exchange is approached.. Increasing number of the corrosion pits is a potential thread to the pipeline condition and the user is obliged to answer the question in which moment the activities aimed at the pipeline replacement into the new one should be undertaken.

The authors of the present study proposed one of possible manners of the problem solution and the described concept should be useful for users of pipelines, in which electro-chemical corrosion is a dominant process of deterioration of the object condition.

References

1. Brodny J.: Wyznaczanie współczynnika dynamicznego przy impulsowym działaniu siły. Prace naukowe GIG. Górnictwo i Środowisko. Wyd. spec. nr 7. Katowice 2008.
2. Broniec Z, Szybka J, Tarnowski J. Modelowanie systemu eksploatacji sieci gazowych użytkowanych na terenach górniczych. Zagadnienia Eksploatacji Maszyn, PWN Warszawa 1998; Z. 1 (113), vol. 33: 65-78.
3. Caleyo F, Gonzalez J. L, Hallen J. M. A study on the reliability assessment methodology for pipelines with active corrosion defects. International Journal of Pressure Vessels and Piping 2002; 79: 77-86.
4. Guian Qian, Markus Niffenegger, Shuxin Li. Probabilistic analysis of pipelines with corrosion defects by using FITNET FSS procedure. Corrosion Science 2011; 53: 855-861.
5. Kucheryavyi V. I, Mil'kov S. N.. Probabilistic prediction of the residual lifetime of a gas pipeline under pitting corrosion. Journal of

Machinery Manufacture and Reliability 2011, vol. 40, no. 5: 489-493.
6. Lenkiewicz W, Szybka J.: Koncepcja modelowania procesów w aspekcie oceny ich niezawodności. Metody prognozowania w inżynierii niezawodności. XXXI Zimowa Szkoła Niezawodności. Szczyrk, 2003.
7. Niezgodziński M. E, Niezgodziński T.: Wzory, wykresy i tablice wytrzymałościowe. WNT, Warszawa 2006.
8. Ponomariew S. D. Współczesne metody obliczeń wytrzymałościowych w budowie maszyn. PWN, Warszawa 1957.
9. Szybka J, Pilch R, Broniec Z, Tarnowski J. Ocena ryzyka eksploatacji ciepłociągu. Problemy Eksploatacji – Zeszyty Naukowe Instytutu Technologii Eksploatacji – PIB nr 1. Radom, 2011; 175-183.
10. Szybka J, Pilch R, Broniec Z. Forecasting the failure of a thermal pipeline on the basis of risk assessment and exploitation analysis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2011; 4(52): 5-10.
11. Szybka J. i inni: Sprawozdanie z badań rurociągu – materiały wewnętrzne niepublikowane. PNTTE, kom. K-ka, 2010.
12. Yong Song, Lin Ma, Jon Morris. A practical approach for reliability prediction of pipeline systems. European Journal of Operational Research 2009; 198: 210-214.
13. PN-80H-74219: Rury bez szwu walcowane na gorąco ze stali węglowej i stopowej do budowy przewodów i konstrukcji.
14. PN-EN ISO 8044: Korozja metali i stopów.

Dr inż. Robert PILCH
Prof. AGH Jan SZYBKA
Dr inż. Zdzisław BRONIEC
AGH University of Science and Technology
Faculty of Mechanical Engineering and Robotics
Department of Machine Design and Technology
Al. A. Mickiewicza 30, 30-059 Cracow, Polad
E-mail: pilch@agh.edu.pl; szybja@agh.edu.pl

Article citation info:
LI YF, HUANG HZ, LIU Y, XIAO N, LI H. A new fault tree analysis method: fuzzy dynamic fault tree analysis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 208-214.

Yan-Feng LI
Hong-Zhong HUANG
Yu LIU
Ningcong XIAO
Haiqing LI

A NEW FAULT TREE ANALYSIS METHOD: FUZZY DYNAMIC FAULT TREE ANALYSIS

NOWA METODA ANALIZY DRZEWA USZKODZEŃ: ROZMYTA ANALIZA DYNAMICZNEGO DRZEWA USZKODZEŃ

Fault tree analysis (FTA) is a widely used reliability assessment tool for large and complex engineering systems. The conventional fault tree analysis method, which contains AND, OR, and Voting gates, etc., can efficiently build an analytical model to represent combinations of component failures that cause the failure of a system. However, due to its limited modeling capability, we may confront difficulties when modeling dynamic systems which involve complicated dynamic characteristics such as sequence dependency and functional dependency. Markov-based dynamic fault tree analysis (DFTA) extends the static FTA by introducing additional gates to model such complicated interactions among events. In many circumstances, it is quite difficult to obtain an accurate system reliability estimate due to limited data. To overcome this issue, a fuzzy dynamic fault tree model is put forth to assess system reliability. To obtain the membership function of the fuzzy probability for the top event of the studied fault trees, the extension principle is employed to calculate the associated membership function via a pair of parametric programming problems. Finally, a case study is presented to demonstrate the application of the proposed approach for the hydraulic system of a CNC machining centre.

***Keywords**: Fault tree analysis, Dynamic fault tree, Fuzzy number, Fuzzy Markov model, Parametric programming.*

Analiza drzewa uszkodzeń (FTA) znajduje szerokie zastosowanie jako narzędzie oceny niezawodności dużych i złożonych systemów inżynierskich. Tradycyjna metoda analizy drzewa uszkodzeń z bramkami logicznymi typu AND, OR, k-z-n, itd. pozwala na sprawne konstruowanie modeli analitycznych reprezentujących kombinacje uszkodzeń elementarnych składowych systemu, które prowadzą do awarii systemu jako całości. Jednakże ograniczone możliwości modelowania jakie daje ta metoda mogą prowadzić do trudności przy modelowaniu systemów dynamicznych posiadających złożone charakterystyki dynamiczne, takie jak zależność sekwencyjna czy zależność funkcjonalna. Analiza dynamicznych drzew uszkodzeń (DFTA) oparta na metodzie Markowa stanowi rozszerzenie tradycyjnej FTA. Wprowadza ona dodatkowe bramki, pozwalając na modelowanie wspomnianych wyżej złożonych interakcji między zdarzeniami. W wielu okolicznościach, ograniczone dane nie pozwalają na otrzymanie dokładnej oceny niezawodności systemu. By rozwiązać ten problem, zaproponowano zastosowanie rozmytego modelu dynamicznego drzewa uszkodzeń do oceny niezawodności systemu. Aby otrzymać funkcję przynależności rozmytego prawdopodobieństwa wystąpienia zdarzenia szczytowego badanego drzewa uszkodzeń, obliczono, na podstawie pary problemów programowania parametrycznego, skojarzoną funkcję przynależności wykorzystując zasadę rozszerzenia. Na zakończenie przedstawiono studium przypadku, w którym proponowane podejście zastosowano do analizy systemu hydraulicznego centrum obróbkowego CNC.

***Słowa kluczowe**: Analiza drzewa uszkodzeń, dynamiczne drzewo uszkodzeń, liczba rozmyta, rozmyty model Markowa, programowanie parametryczne.*

Introduction

Fault tree analysis (FTA) is a logical and graphic method being widely used to evaluate the reliability of complex engineering systems from both qualitative and quantitative perspectives. Fault tree provides a graphical representation of combinations of component failures leading to an undesired system failure [22, 25].

Fault tree (FT) was first introduced in 1961 by H. A. Watson of Bell telephone laboratories in connection with a U.S. Air Force contract to study the minuteman missile launch control system [5]. In the 1965 safety symposium, sponsored by the University of Washington and the Boeing company, several papers were presented, which expounded the virtues of fault-tree analysis [21]. Since then, a variety of methods for modeling and evaluating the complex system reliability via FTA have been reported [1, 2, 4, 6].

However, in many situations, the behaviour of components in a complex system and their interactions, such as failure priority, sequentially dependent failures, functional dependent failures, and dynamic redundancy management, cannot be adequately addressed by traditional FT due to its limited modeling capability. Several approaches have been proposed to overcome these difficulties. Dugan et al. [9, 10, 16] introduced a modularization method to identify the independent sub-trees with dynamic gates. Markov models were used to solve these dynamic fault trees. Amari et al. [3] proposed a numerical integration technique to solve dynamic gates without converting them to Markov models. By using the probability distribution and conditional probability distribution of the basic events, this method can accurately assess the fault tree behaviour with dynamic gates and repeated basic events. Bobbio et al. [7-8] proposed a Bayesian network-based method to further reduce the complexity of solving dynamic FTs based on a state-space approach.

The aforementioned state-based approaches are capable of evaluating the reliability of complex systems either qualitatively or quan-

titatively. We can obtain the accurate failure probability value of the top event by using these state-based fault tree analysis methods. When performing a system reliability evaluation, both operation states and failure states of components are generally assumed to be known. That is to say, components are assumed in either functioning state or failed state, and the probability for being in each individual state can be determined in advance. Actually, this is not always the case in many real engineering systems due to two main reasons:

(1) The states of components/systems often deteriorate over time, so failures of these components/systems may not occur at a certain point in time. Sometimes, we cannot exactly identify the state of a component or a system due to various kinds of uncertain factors, such as inaccurate measurements and human errors. In addition, ambiguity of system and component behaviour, and the dynamic operating environment of a system introduces additional difficulties in estimating the exact failure probabilities of basic events.

(2) Obtaining large and accurate failure data is costly, difficult, or even impossible for many real and complex systems. This is true especially for those systems with components whose failure rates are very low and/or with new designs. In these situations, it is not realistic or possible to represent the component failure behaviour using crisp values.

Fuzzy set theory proposed by Zadeh [23-24] has shown to be a useful methodology to cope with these cases where subjective judgement or estimation of an individual plays a vital and useful role in dealing with the ambiguity or uncertainty. Many papers have been published to incorporate the fuzzy set theory into the fault tree analysis for reliability analysis. Tanaka et al. [20] proposed an enumeration approach to estimate the cut sets of FT for which the trapezoidal fuzzy number is used to represent failure probabilities of events. Singer [19] used triangular fuzzy numbers to substitute the exact probability value as a representation of failure probabilities for basic events and top events. Based on the extension principle, Misra and Weber [17], Liang and Wang [13], described the fuzzy arithmetic operations for fault tree analysis. To avoid uncertainty in probabilistic risk assessment, Singer [19], Lai et al. [12] and Sawyer [18] introduced fuzzy set theory into safety and reliability modeling process. Based on posbist reliability theory, posbist fault tree analysis of coherent systems was proposed by Huang et al. [11]. In their approach, event failure behaviour is characterized in the context of possibilistic measures, and the structure function of the posbist fault tree of a coherent system is defined.

In this paper, triangular fuzzy membership functions are used to describe the vagueness of quantification of failure probability for basic events. Based on a fuzzy transition rate matrix, a fuzzy Markov model is introduced to capture dynamic behaviour of systems. Finally, a numerical example is provided to illustrate the application of the proposed method.

2. Dynamic fault tree

2.1. Dynamic gates

A major disadvantage of the traditional FTA is its inability to capture sequence dependencies in the system while still allowing an analytic solution [9-10]. To overcome this difficulty, Dugan proposed a new reliability analysis method called Dynamic Fault Tree Analysis (DFTA) by introducing several dynamic gates to describe the dynamic behaviours of these systems [9, 10, 16]. There are four major basic dynamic gates which will be elaborated in follows.

2.1.1. Priority-AND gate (PAND gate)

The PAND gate has two inputs, A and B, both of which could be a basic event or the output of other logic gates. The PAND gate reaches a failure state if all its input components have failed in a pre-assigned order (generally from left to right in a graphical notation).

2.1.2. Functional dependency gate (FDEP gate)

The FDEP gate frequently includes one trigger input (either a basic event or a output of another gate) and one or more dependent events. The dependent events are functionally relying on the trigger event. When the trigger event occurs, the dependent basic events are forced to occur.

2.1.3. Sequence enforcing gate (SEQ gate)

The SEQ gate forces its inputs to fail in a particular order. It never happens if the failure sequence takes place in different orders. While the SEQ gate allows events to occur only in a pre-specified order and state that a different failure sequence can never take place, the PAND gate does not impose such a strong assumption: it simply detects the failure order and a failure triggered upon the match with the order.

2.1.4. Spare gates

The spares often include one principal component that can be substituted by one or more backups that have the same function with the principal one. The Spare gate is classified into three types of backups, i.e., Cold Spare (CSP), Warm Spare (WSP) and Hot Spare (HSP). Suppose λ being the failure rate of a component, and then the failure rate alters to $\alpha\lambda$ while being used as a spare. If $\alpha = 0$, the spare is a CSP and $\alpha = 1$, the spare is a HSP, otherwise it's a WSP for the case where $0<\alpha<1$.

The four different categories of dynamic gates are enumerated in Table 1 with their input information, failure criteria, and corresponding symbols.

2.2. Markov model

In a dynamic fault tree, the occurrence of a top event depends not only on the combination of component failures, but also on the sequence of occurrences of these events. Thus, the Markov model has been used as a quantitative method to model the failure process and evaluate system reliability.

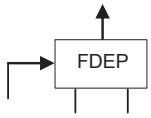
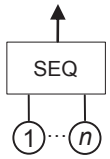
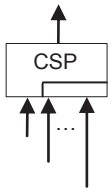
Let T be an infinite real set, $t \in T$, then $\{X(t), t \in T\}$ is called a stochastic process. For each $t \in T$, $X(t)$ is a random variable. A typical continuous stochastic process $\{X(t), t \in T\}$ is called a Markov process if its conditional probability satisfies the relation

$$\begin{aligned} P\{X(t_n) = x_n | X(t_1) = x_1, X(t_2) = x_2, \cdots, X(t_{n-1}) = x_{n-1}\} \\ = P\{X(t_n) = x_n | X(t_{n-1}) = x_{n-1}\} \end{aligned} \tag{1}$$

where $x_i \in S$, S is the state space of the stochastic process and $t_1 < t_2 < \cdots < t_{n-1} < t_n$. This memory-less characteristic means that the probability of this stochastic process being in state x_i at time t_i depends only on the state at time t_{i-1} and is independent of the state at time $t_i (i = 1, 2, \cdots, n-2)$.

The occurrence of component failure is a frequent stochastic process which can be represented by certain types of probability distribution functions. In a dynamic system, the failure process of the system can be represented by a Markov process. Suppose that the system has n states $s_i (i = 1, 2, \cdots, n)$, $s_i \in S$, and S is the state space of the Markov process $\{S(t),\ t \geq 0\}$. The transition rate from state i to state j is denoted by $\lambda_{i,j}$. Then the failure process of the system can be represented by a transition diagram shown in Fig. 1.

Table 1. Dynamic gates and failure mechanism

Dynamic gate	Input events information	Failure criteria	Symbol
PAND Gate	The PAND Gate has two inputs, A and B, both of which can be basic events or outputs of other logic gates.	The output of this gate is true if both inputs have occurred, and A occurred before B.	
FDEP Gate	The FDEP Gate has a trigger event and multiple dependent basic events.	If the trigger event occurs, all the dependent events occur subsequently, and the output becomes true.	
SEQ Gate	The SEQ Gate has multiple inputs	The sequence-enforcing gate (SEQ) forces its inputs to occur in a particular order. If all the inputs occur, the output is true.	
Spare Gate	The Spare has one primary input and one or more alternate inputs	If the primary unit fails, the first alternate component begins to function, till all the replacements fail, the output becomes true.	

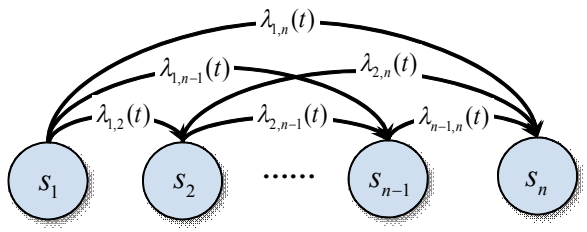


Fig. 1. A sample state transition diagram

Let $P_i(t),\ i=1,2,\cdots,n$ be the probability of the system being on state $s_i(i=1,2,\cdots,n)$, the differential equations for the aforementioned dynamic process take the form as follows:

$$\begin{cases} \dfrac{dp_1(t)}{dt} = -p_1(t)\sum_{j=2}^{n}\lambda_{1,j} \\ \dfrac{dp_i(t)}{dt} = \sum_{j=1}^{i-1} p_j(t)\lambda_{j,i} - \sum_{j=i+1}^{n} p_i(t)\lambda_{i,j},\ \ 1<i<n,\ t\ge 0 \\ \dfrac{dp_n(t)}{dt} = \sum_{j=1}^{n-1} p_j(t)\lambda_{j,n} \end{cases} \tag{2}$$

Solving the equations with the initial condition:

$$\begin{cases} P_1(0)=1, \\ P_i(0)=0, i=2,\cdots,n \end{cases}$$

we can obtain the probability value $P_n(t)$, which is also the failure probability of the top event corresponding to the Markov state transition diagram.

3. Fuzzy set theory

3.1. Fuzzy set and fuzzy number

During the evaluation of reliability of complex engineering systems, there exist two kinds of uncertainties, i.e. aleatory uncertainty and epistemic uncertainty. Zadeh [23-24] proposed a set of systematic mathematical theories, namely fuzzy set theory, which can deal with fuzzy characteristics of uncertainty in real engineering systems.

Given a universal set U, for a set $\tilde{A}$, and for each u, there exits a real number $\mu_{\tilde{A}}(u)\in[0,1]$ that corresponds to u, which represents the degree of u belonging to $\tilde{A}$. We call the set $\tilde{A}$ a fuzzy set, and the value $\mu_{\tilde{A}}(u)$ the membership degree of u to $\tilde{A}$.

$$\begin{aligned} \mu_{\tilde{A}}:\ & U\to[0,1] \\ & u \mapsto \mu_{\tilde{A}}(u) \end{aligned}$$

For a fuzzy set $\tilde{A}$, it becomes a fuzzy number if it is a normal as well as a convex fuzzy set. Triangular fuzzy number, normal fuzzy number, and trapezoidal fuzzy number are among the mostly used fuzzy numbers.

A typical triangular fuzzy number $\tilde{A}(a,b,c)$ can be defined by its membership function as follows. Its graphic representation is shown in Fig. 2.

$$\mu_{\tilde{A}} = \begin{cases} \dfrac{x-a}{b-a}, & a\le x<b \\ 1, & x=b \\ \dfrac{x-c}{b-c}, & b<x\le c \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

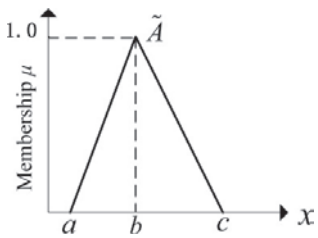


Fig. 2. Membership function of the triangular fuzzy number

3.2. Extension principle

The concept of fuzzy set and fuzzy number proposed by Zadeh [23-24] provides a means to represent and quantify fuzzy information. Besides, Zadeh introduced the extension principle for fuzzy operations between fuzzy numbers.

Given $\tilde{X}_i(i=1,2,\cdots,n)$ are fuzzy numbers corresponding to universal set $R_i(i=1,2,\cdots,n)$, respectively. $x_i\in R_i(i=1,2,\cdots,n)$ are the variables associated with each fuzzy number $\tilde{X}_i(i=1,2,\cdots,n)$. $f(x_1, x_2,...,x_n)$ is a function that maps the variables $x_i(i=1,2,\cdots,n)$ to a variable $y(y\in R)$. Then we can induce (generate) a fuzzy number

$\tilde{Y}$ from fuzzy numbers $\tilde{X}_i (i=1,2,\cdots,n)$ by function $f(x_1, x_2,...,x_n)$. The membership function of $\tilde{Y}$ can be obtained by the extension principle shown as follow:

$$u_{\tilde{Y}}(y) = \sup_{\substack{x_i \in R_i (i=1,2,\cdots,n) \\ y=f(x_1,x_2,\cdots,x_n)}} \min(u_{\tilde{X}_1}(x_1), u_{\tilde{X}_2}(x_2), \cdots, u_{\tilde{X}_n}(x_n)) \tag{4}$$

According to the extension principle, the interval of the α-cut of the fuzzy number $\tilde{Y}$ is given by:

$$\begin{aligned} \tilde{Y}_\alpha(y) &= [\min_{1\le i\le n} f(x;\mu_{\tilde{x}_i}(x_i)\ge\alpha), \max_{1\le i\le n} f(x;\mu_{\tilde{x}_i}(x_i)\ge\alpha)] \\ &= [\tilde{Y}_\alpha{}^L, \tilde{Y}_\alpha{}^U] \end{aligned} \tag{5}$$

Thus, the lower and upper bounds of $\tilde{Y}$ can be obtained by following a pair of parametric programming problems

$$\begin{aligned} &\tilde{Y}_{\mathrm{a}}^L = \min f(x_1, x_2, \cdots, x_n) \\ &\text{subject to} \\ &\tilde{x}_{1\mathrm{a}}^L \le x_1 \le \tilde{x}_{1\mathrm{a}}^U \\ &\tilde{x}_{2\mathrm{a}}^L \le x_2 \le \tilde{x}_{2\mathrm{a}}^U \\ &\quad\vdots \\ &\tilde{x}_{n\mathrm{a}}^L \le x_n \le \tilde{x}_{n\mathrm{a}}^U \end{aligned} \tag{6}$$

$$\begin{aligned} &\tilde{Y}_{\mathrm{a}}^U = \max f(x_1, x_2, \cdots, x_n) \\ &\text{subject to} \\ &\tilde{x}_{1\mathrm{a}}^L \le x_1 \le \tilde{x}_{1\mathrm{a}}^U \\ &\tilde{x}_{2\mathrm{a}}^L \le x_2 \le \tilde{x}_{2\mathrm{a}}^U \\ &\quad\vdots \\ &\tilde{x}_{n\mathrm{a}}^L \le x_n \le \tilde{x}_{n\mathrm{a}}^U \end{aligned} \tag{7}$$

We can easily obtain the intervals at different α-cut levels by the extension principle.

4. Fuzzy dynamic fault tree (FDFT)

Combining the Markov model with the fuzzy set theory, we propose a reliability analysis method called Fuzzy Dynamic Fault Tree (FDFT) to estimate the reliability of systems having dynamic characteristics and fuzzy uncertainty.

The fuzzy Markov model corresponds to the Dynamic Fault Tree (DFT) with n states $s_i, 1 \le i \le n$, and the crisp state transition rate $\lambda_{i,j}$ of Markov model is replaced by fuzzy state transition rate $\tilde{\lambda}_{i,j}$ due to difficulty of estimating accurate values. Then the fuzzy state transition rate matrix $\tilde{A}$ is given as follows:

$$\tilde{A} = \left(\tilde{\lambda}_{i,j}\right) = \begin{pmatrix} \tilde{\lambda}_{1,1} & \tilde{\lambda}_{1,2} & \cdots & \tilde{\lambda}_{1,n} \\ \tilde{\lambda}_{2,1} & \tilde{\lambda}_{2,2} & \cdots & \tilde{\lambda}_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{\lambda}_{n,1} & \tilde{\lambda}_{n,2} & \cdots & \tilde{\lambda}_{n,n} \end{pmatrix} \tag{8}$$

Thus, the state transition diagram is given as Fig. 3.

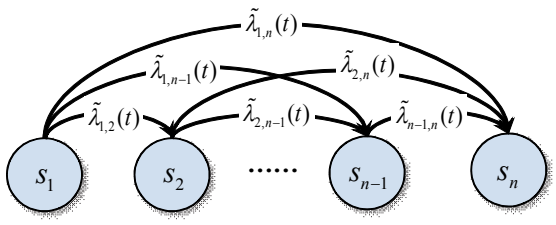


Fig. 3 The fuzzy state transition diagram for a non-repairable system

As a result, the differential equations with the fuzzy transition rate take the form of:

$$\begin{cases} \dfrac{d\tilde{p}_1(t)}{dt} = -\tilde{p}_1(t)\sum_{j=2}^{n}\tilde{\lambda}_{1,j} \\ \dfrac{d\tilde{p}_i(t)}{dt} = \sum_{j=1}^{i-1}\tilde{p}_j(t)\tilde{\lambda}_{j,i} - \sum_{j=i+1}^{n}\tilde{p}_i(t)\tilde{\lambda}_{i,j}, \quad 1<i<n, \ t\ge 0 \\ \dfrac{d\tilde{p}_n(t)}{dt} = \sum_{j=1}^{n-1}\tilde{p}_j(t)\tilde{\lambda}_{j,n} \end{cases} \tag{9}$$

To solve the differential equations, the Laplace-Stieltjes transform can be used with the initial condition: $\tilde{p}_1(0)=1, \tilde{p}_i(0)=0 (i\ne 1)$.

The corresponding linear equations take the form as follows:

$$\begin{cases} s\tilde{p}_1(s) - 1 = -\tilde{p}_1(s)\sum_{i=2}^{n}\tilde{\lambda}_{1,i} \\ s\tilde{p}_i(s) = \sum_{j=1}^{i-1}\tilde{p}_j(s)\tilde{\lambda}_{j,i} - \sum_{j=i+1}^{n}\tilde{p}_i(s)\tilde{\lambda}_{i,j}, \quad 1<i<n \\ s\tilde{p}_n(s) = \sum_{j=1}^{n-1}\tilde{p}_j(s)\tilde{\lambda}_{j,n} \end{cases} \tag{10}$$

After solving the linear equations to obtain $\tilde{p}_n(s)$, and the inverse Laplace-Stieltjes transform can be used to solve $\tilde{p}_n(t)$. Then the interval of the probability of $\tilde{p}_n(t)$ can be solved by the extension principle mentioned in Section 3.

5. Reliability analysis for CNC via FDFT

5.1. Brief introduction of CNC hydraulic systems

A complete hydraulic system is mainly composed of five parts: power components, control components, executive components, auxiliary components, and hydraulic oil. The studied system here is composed of four circuits, i.e. spindle balancing circuit, spindle releasing circuit, C axle clamping and releasing circuit, and D axle clamping and releasing circuit. In this section the Spindle balancing circuit is analysed by using FDFT model. The configuration of the spindle is shown in Fig. 4. It is composed of a tank, three filters, one hydraulic gear pump, three shut-off valves, check valve, one pressure relief valve, two pressure gauges, one pressure relay, two overflow valves, one cylinder, and one power accumulator.

Its operation principle can be described as follows. The oil is pumped from the tank to the main oil line through the control valves aforementioned. When the pressure value exceeds 140 bar, the oil will flow back to the tank through the overflow valve to reduce the pressure down to 140 bar. The pressure will drop to 65 bar through the pressure relief valve. The power accumulator can be accumulated when the pressure is over 50 bar. The operation of the pump will be blackout by the pressure delay when the pressure of the main circuit is over 60 bar. Afterwards, it will start again until the pressure value drops below 55 bar.

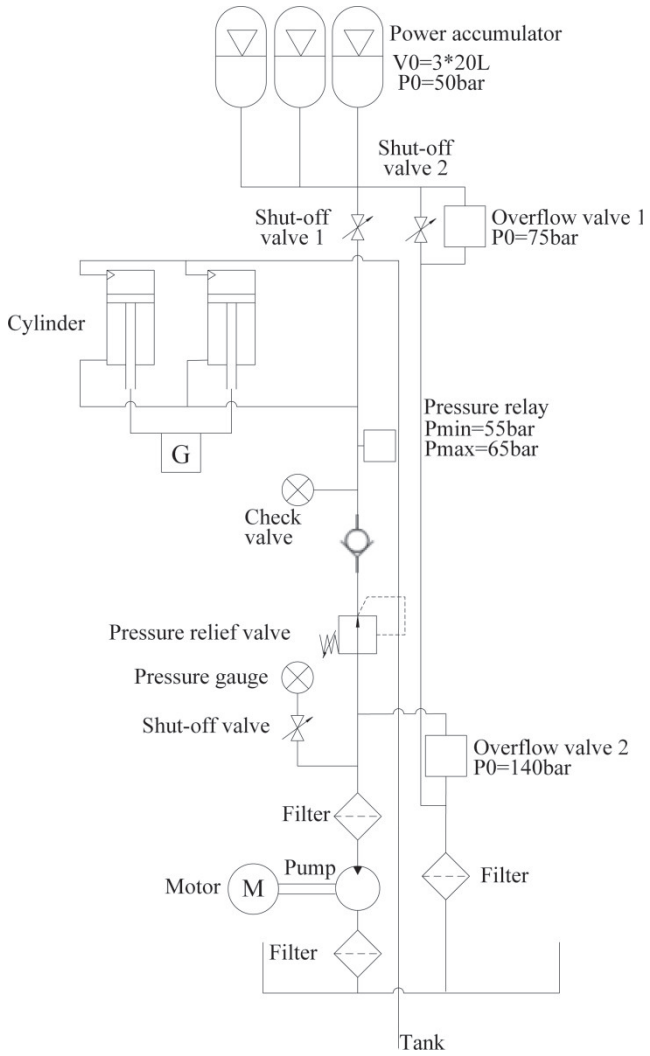


Fig. 4 Spindle carrier balance circuit of CNC machining center

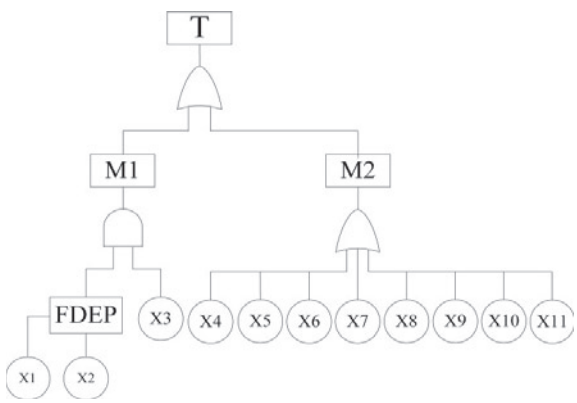


Fig. 5 Dynamic fault tree of the hydraulic system

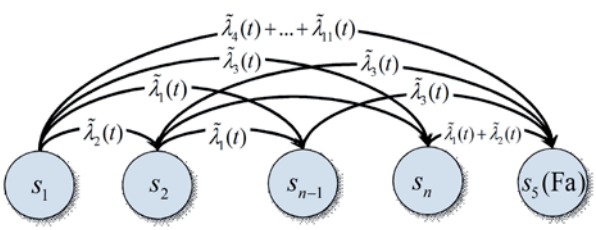


Fig. 6 State transition diagram of the spindle circuit of hydraulic system`

5.2. Construction of dynamic fault tree of spindle balance circuit

The event of the insufficiency of pressure in the balance circuit is considered as the top event in the following analysis. The filter failure and pressure gauge failure are ignored due to their low failure rate. The basic events are enumerated as follows.

x_1: pressure relay failure; x_2: pump failure; x_3: power accumulator failure; x_4: tank failure; x_5: check valve failure; x_6: cylinder failure; x_7: shut-off valve 1 failure; x_8: shut-off valve 2 failure; x_9: pressure relief valve failure; x_{10}: overflow valve 1 failure; x_{11}: overflow valve 2 failure. The dynamic fault tree is shown in Fig. 5.

5.3. Quantitative assessment of FDFT

According to the FDFT shown in Section 5.2 and the basic failure data shown in Table 2, the reliability of the spindle in hydraulic system is analysed based on FDFT. The failure rate of each component is represented by triangular fuzzy numbers shown in Table 2. The DFT is transformed into the fuzzy Markov model shown in Fig. 6.

The corresponding fuzzy state transition rate matrix is as follows:

$$\tilde{A}=\begin{pmatrix} -\sum_{i=1}^{11}\tilde{\lambda}_i & \tilde{\lambda}_2 & \tilde{\lambda}_1 & \tilde{\lambda}_3 & \sum_{i=4}^{11}\tilde{\lambda}_i \\ 0 & -\tilde{\lambda}_1-\tilde{\lambda}_3 & \tilde{\lambda}_1 & 0 & \tilde{\lambda}_3 \\ 0 & 0 & -\tilde{\lambda}_3 & 0 & \tilde{\lambda}_3 \\ 0 & 0 & 0 & -\tilde{\lambda}_1-\tilde{\lambda}_2 & \tilde{\lambda}_1+\tilde{\lambda}_2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The dynamic differential equations take the form:

$$\begin{cases} \dfrac{d\tilde{p}_1(t)}{dt}=-\tilde{p}_1(t)\sum_{i=1}^{11}\tilde{\lambda}_i \\ \dfrac{d\tilde{p}_2(t)}{dt}=\tilde{p}_1(t)\tilde{\lambda}_2-\tilde{p}_2(t)(\tilde{\lambda}_1+\tilde{\lambda}_3) \\ \dfrac{d\tilde{p}_3(t)}{dt}=\tilde{p}_1(t)\tilde{\lambda}_1+\tilde{p}_2(t)\tilde{\lambda}_1-\tilde{p}_3(t)\tilde{\lambda}_3 \\ \dfrac{d\tilde{p}_4(t)}{dt}=\tilde{p}_1(t)\tilde{\lambda}_3-\tilde{p}_4(t)(\tilde{\lambda}_1+\tilde{\lambda}_2) \\ \dfrac{d\tilde{p}_5(t)}{dt}=\tilde{p}_1(t)\sum_{i=4}^{11}\tilde{\lambda}_i+(\tilde{p}_2(t)+\tilde{p}_3(t))\tilde{\lambda}_3+\tilde{p}_4(t)(\tilde{\lambda}_1+\tilde{\lambda}_2) \end{cases} \qquad t\geq 0$$

Table 2. Basic event and failure rate

Basic event	Fuzzy failure rate $\tilde{\lambda}(t)$ (×10^{-6})	Basic event	Fuzzy failure rate $\tilde{\lambda}(t)$ (×10^{-6})
x_1	(0.0425,0.0500,0.0575)	x_7	(7.2250,8.5000,9.7750)
x_2	(11.4750,13.5000,15.5250)	x_8	(7.2250,8.5000,9.7750)
x_3	(6.1200,7.2000,8.2800)	x_9	(1.8190,2.1400,2.4610)
x_4	(1.2750,1.5000,1.7250)	x_{10}	(4.8450,5.7000,6.5550)
x_5	(4.2500,5.0000,5.7500)	x_{11}	(4.8450,5.7000,6.5550)
x_6	(0.0068,0.0080,0.0092)		

with the initial conditions $\tilde{p}_1(0)=1$ and $\tilde{p}_i(0)=0\ (1<i\le 5)$. The Laplace-Stieltjes transform can be used to solve the differential equations. After the transform, the differential equations become a set of linear equations as follows:

$$\begin{cases} s\tilde{p}_1(s)-1=-\tilde{p}_1(s)\sum_{i=1}^{11}\tilde{\lambda}_i \\ s\tilde{p}_2(s)=\tilde{p}_1(s)\tilde{\lambda}_2-\tilde{p}_2(s)(\tilde{\lambda}_1+\tilde{\lambda}_3) \\ s\tilde{p}_3(s)=\tilde{p}_1(s)\tilde{\lambda}_1+\tilde{p}_2(s)\tilde{\lambda}_1-\tilde{p}_3(s)\tilde{\lambda}_3 \\ s\tilde{p}_4(s)=\tilde{p}_1(s)\tilde{\lambda}_3-\tilde{p}_4(s)(\tilde{\lambda}_1+\tilde{\lambda}_2) \\ s\tilde{p}_5(s)=\tilde{p}_1(s)\sum_{i=4}^{11}\tilde{\lambda}_i+(\tilde{p}_2(s)+\tilde{p}_3(s))\tilde{\lambda}_3+\tilde{p}_4(s)(\tilde{\lambda}_1+\tilde{\lambda}_2) \end{cases}$$

Then, the fuzzy probability of state s_5 can be obtained as

$$\tilde{p}_5(s)=\frac{1}{s}-\frac{\tilde{\lambda}_3}{(\tilde{\lambda}_3+\sum_{i=4}^{11}\tilde{\lambda}_i)\times(s+\tilde{\lambda}_1+\tilde{\lambda}_2)}-\frac{\tilde{\lambda}_1+\tilde{\lambda}_2}{(\tilde{\lambda}_1+\tilde{\lambda}_2+\sum_{i=4}^{11}\tilde{\lambda}_i)\times(s+\tilde{\lambda}_3)}+\frac{(\tilde{\lambda}_1\times\tilde{\lambda}_3+\tilde{\lambda}_2\times\tilde{\lambda}_3-(\sum_{i=4}^{11}\tilde{\lambda}_i)^2}{(\tilde{\lambda}_3+\sum_{i=4}^{11}\tilde{\lambda}_i)\times(\tilde{\lambda}_1+\tilde{\lambda}_2+\sum_{i=4}^{11}\tilde{\lambda}_i)\times(s+\sum_{i=1}^{11}\tilde{\lambda}_i)}$$

By taking the inverse Laplace-Stieltjes transform, the fuzzy probability of state s_5 can be derived as a function of t.

$$\tilde{p}_5(t)=1-\frac{\tilde{\lambda}_3}{\tilde{\lambda}_3+\sum_{i=4}^{11}\tilde{\lambda}_i}\times\exp(-(\tilde{\lambda}_1+\tilde{\lambda}_2)\times t)-\frac{\tilde{\lambda}_1+\tilde{\lambda}_2}{\tilde{\lambda}_1+\tilde{\lambda}_2+\sum_{i=4}^{11}\tilde{\lambda}_i}\times\exp(-\tilde{\lambda}_3\times t)+\frac{(\tilde{\lambda}_1\times\tilde{\lambda}_3+\tilde{\lambda}_2\times\tilde{\lambda}_3-(\sum_{i=4}^{11}\tilde{\lambda}_i)^2)\times\exp(-(\tilde{\lambda}_1+\tilde{\lambda}_2+\tilde{\lambda}_3+\sum_{i=4}^{11}\tilde{\lambda}_i)\times t)}{(\tilde{\lambda}_3+\sum_{i=4}^{11}\tilde{\lambda}_i)\times(\tilde{\lambda}_1+\tilde{\lambda}_2+\sum_{i=4}^{11}\tilde{\lambda}_i)}$$

Using Eqs. (4)-(7), the membership of fuzzy probability of state s_5 at different cut level α can be calculated as shown in Figs.7-9.

The membership function of fuzzy failure probability presented in Fig.7 is obtained by Eqs. (6) and (7) at different α-cut level at time $t=5000$ hours. The median value of fuzzy failure probability is 0.1631 at time $t=5000$ hours. The membership functions of fuzzy failure probability at time $t=10000$ and $t=15000$ are obtained by the same way with the median value 0.2892 and 0.3875 shown in Fig.8 and Fig.9

The reliability curve is presented in Fig.10. The blue one is obtained by the fuzzy failure rates corresponding to α =0. The red one which falls into the blue one is obtained by the crisp failure rates corresponding to α =1.

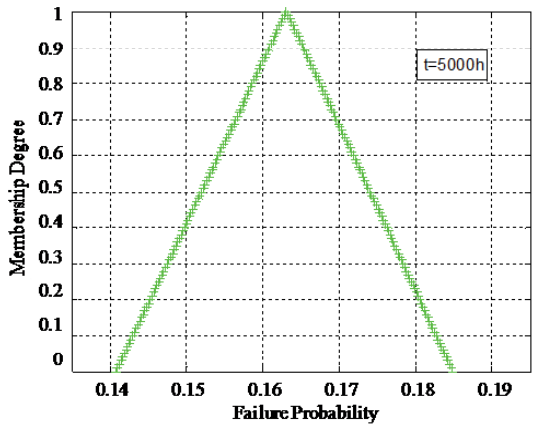


Fig.7 The membership of $\tilde{p}_5$ at $t=5000$ h

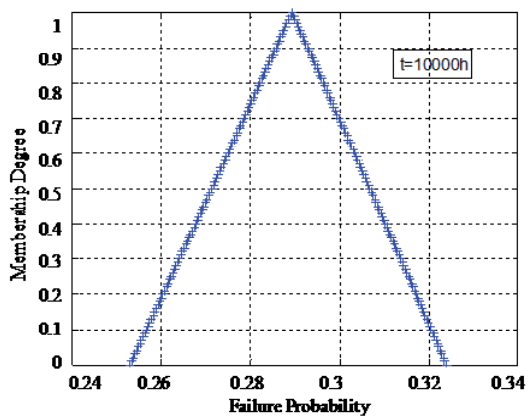


Fig.8 The membership of $\tilde{p}_5$ at $t=10000$ h

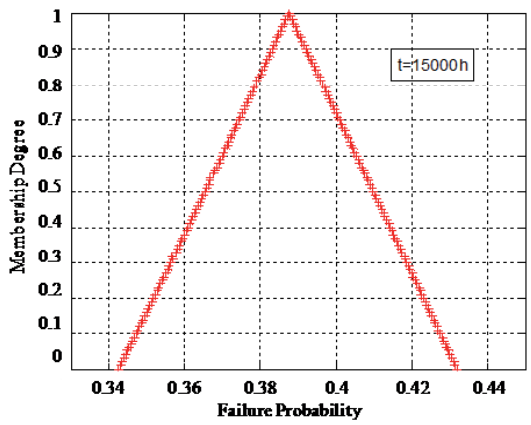


Fig.9 The membership of $\tilde{p}_5$ at $t=15000$ h

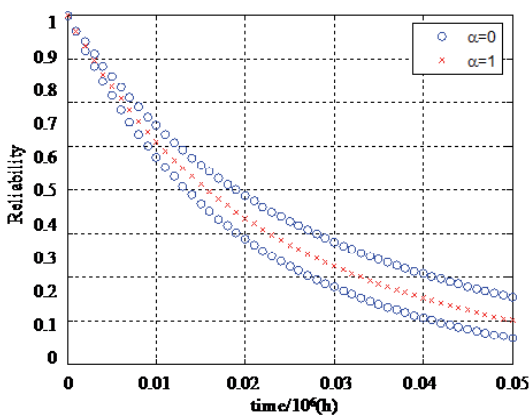


Fig.10 The fuzzy reliability of the system with α = 0 and α = 1

6. Conclusion

This paper applied the fuzzy Markov model to evaluate the reliability of a complex mechanical system used in CNC machine centre. When modeling of the fault tree, the dynamic logic gates are used to capture the dynamic behaviour in the system. The Markov model and Laplace-Stieltjes transform are used to solve the fault tree model. Fuzzy set theory is shown to be quite effective in quantifying the information uncertainty in the system under study. Triangular fuzzy numbers and the extension principle are used to solve the Markov model with fuzzy uncertainty. The result shows that the proposed method is a promising approach to reliability analysis.

Acknowledgement

This research was partially supported by Important National Science & Technology Specific Projects of China under the contract number 2009ZX04002-013 and Specialized Research Fund for the Doctoral Program of Higher Education of China under the contract number 20090185110019.

7. References

1. Aggarwal K K. Comment on an efficient simple algorithm for fault tree automatic synthesis from the reliability graph. IEEE Trans. Reliability 1979; 28(4): 309.
2. Allan R N, Rondiris I L, Fryer D M. An efficient computational technique for evaluating the cut/tie sets and common-cause failures of complex systems. IEEE Trans. Reliability 1981; 30(2): 101-109.
3. Amari S, Dill G, Howald E. A new approach to solve dynamic fault trees. Proceedings of Annual IEEE Reliability and Maintainability symposium 2003; 374-379.
4. Andow P K. Difficulties in fault-tree synthesis for process plant. IEEE Trans. Reliability 1980; 29(1): 2-9.
5. Bell Telephone Laboratories. Launch control safety study, Section VII. Murray Hill Press: N.J. USA. 1961.
6. Bengiamin N N, Bowman B A, Schenk K F. An efficient algorithm for reducing the complexity of computation in fault tree analysis. IEEE Trans. Nuclear Science 1976; 23(5): 1442-1446.
7. Bobbio A, Daniele C R. Parametric fault trees with dynamic gates and repair boxes. Proceedings of the Annual IEEE Reliability and Maintainability Symposium 2004; 459-465.
8. Bobbio A, Portinale L, Minichino M, Ciancamerla E. Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. Reliab Eng Syst Saf 2001; 71: 249-260.
9. Dugan J B, Bavuso S J, Boyd M A. Dynamic fault-tree for fault-tolerant computer systems. IEEE Trans. Reliability 1992; 41(3):363-376.
10. Dugan J B, Sullivan K J, Coppit D. Developing a low cost high-quality software tool for dynamic fault-tree analysis. IEEE Trans. Reliability 2000; 49(1): 49-59.
11. Huang H Z, Tong X, Zuo M J. Posbist fault tree analysis of coherent systems. Reliab Engng Syst Saf 2004; 84(2): 141-148.
12. Lai F S, Shenoi S, Fan T L. Fuzzy Fault Tree Analysis: Theory and Applications. Engineering Risk and Hazard Assessment. CRC Press, 1988.
13. Liang G, Wang J M. Fuzzy fault tree analysis using failure possibility. Microelectronics and Reliability 1993; 33(4): 583-597.
14. Liu Y, Huang H Z. Reliability assessment for fuzzy multi-state systems. International Journal of Systems Science 2010; 41(4): 365-379.
15. Liu Y, Huang H Z, Levitin G. Reliability and performance assessment for fuzzy multi-state elements. Journal of Risk and Reliability 2008; 222(4): 675-686.
16. Meshkat L, Dugan J B, Andrews J D. Dependability analysis of systems with on - demand and active failure modes using dynamic fault trees. IEEE Trans. Reliability 2002; 51(2): 240-251.
17. Misra B K, Weber G G. Use of fuzzy set theory for level-1 studies in probabilistic risk assessment. Fuzzy Sets and System 1990; 37(2): 139-160.
18. Sawyer P J, Rao S S. Fault tree analysis of fuzzy mechanical system. Microelectronics and Reliability 1994; 34(4): 653-667.
19. Singer D. A fuzzy set approach to fault tree and reliability analysis. Fuzzy Sets and Systems 1990; 34(2): 145-155.
20. Tanaka H, Fan L T, Lai F S, Toguchi K. Fault-tree analysis by fuzzy probability. IEEE Trans. Reliability 1983; 32(5): 453-457.
21. The Boeing Company. System Safety Symposium, Available from University of Washington Library. Seattle, Washington, 1965.
22. Vesely W E. Fault Tree Handbook. NUREG-0492. US Nuclear Regulatory Commission: Washington D.C., 1981.
23. Zadeh L A. Fuzzy sets. Information and control 1965; 8: 338-353.
24. Zadeh L A. Fuzzy set as a basis for a theory of possibility. Fuzzy Sets and Systems 1978; 1(1):3-28.
25. Fault Tree Analysis. Edition 2.0. International Electrotechnical Commission (IEC 61025), 2006.

Yan-Feng LI, Ph.D. candidate
Prof. Hong-Zhong HUANG, Ph.D.
Associate Prof. Yu LIU, Ph.D.
Ningcong XIAO, Ph.D. candidate
Associate Prof. Haiqing LI, 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:
ŚWIĆ A, TARANENKO W. Adaptive control of machining accuracy of axial-symmetrical low-rigidity parts in elastic-deformable state. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 215-221.

Antoni ŚWIĆ
Wiktor TARANENKO

ADAPTIVE CONTROL OF MACHINING ACCURACY OF AXIAL-SYMMETRICAL LOW-RIGIDITY PARTS IN ELASTIC-DEFORMABLE STATE

STEROWANIE ADAPTACYJNE DOKŁADNOŚCIĄ OBRÓBKI CZĘŚCI OSIOWO-SYMETRYCZNYCH O MAŁEJ SZTYWNOŚCI W STANIE SPRĘŻYŚCIE-ODKSZTAŁCALNYM*

The authors developed a method of correction consisting in the introduction, in the control system, of an additional positive feedback relative to the force of milling. Adaptive control was applied for axial feed and for additional force actions causing the elastic-deformable state, which permits the elimination of static errors of control effects and interference in the control of quality parameters.

***Keywords**: adaptive control, machining accurcy, low-rigidity parts,elastic-deformable state.*

Opracowano sposób korekty ustawienia układu technologicznego obrabiarki polegający na wprowadzeniu do układu sterowania dodatkowego dodatniego sprzężenia zwrotnego względem siły skrawania. Zastosowane sterowanie adaptacyjne posuwem wzdłużnym oraz dodatkowymi oddziaływaniami siłowymi, wywołującymi stan sprężyście-odkształcalny, umożliwia wyeliminowanie błędów statycznych oddziaływania sterującego oraz zakłócającego przy sterowaniu parametrami dokładności obróbki części o małej sztywności w stanie sprężyście-odkształcalnym.

***Słowa kluczowe**: sterowanie adaptacyjne, dokładność obróbki, części o małej sztywności, stan sprężyście-odkształcalny.*

1. Introduction

In the machine-building industry, axisymmetrical parts constitute approximately 34% of the total number of parts, and among those up to 12% can be classified as low-rigidity shafts.

Such parts are characterised by disproportional relations of overall dimensions and by low rigidity in specific sections and directions. Stringent requirements are also posed relative to the parameters of geometric form, mutual positioning of surfaces, linear dimensions, and surface quality.

The specific character of machining of similar parts causes that it is difficult to achieve the required parameters of accuracy of form, dimensions and surface quality. The low inherent stiffness and the relatively low rigidity of the shaft compared to the stiff assemblies of the machine tool cause the appearance of vibrations, under specific conditions. In the course of machining there appear numerous factors that interfere with and destabilise the process of machining (large free deformations of shafts, tools, fixtures, shavings, dust, etc.). This makes it necessary to search for new methods of controlling the machining of low-rigidity shafts.

Positive solution of problems elated with improvement of quality of machining of parts of milling machines for metals depends directly on further improvement of methods for designing and testing of automatic control systems (ACS), machine tools and parameters of the dynamic system. The dynamic system (DS) of the process of machining is an MGFT (machine tool-grip-fixture-tool) system – it comprises the mass-dissipative-elastic (MDE) system of the machine tool and the realized process of machining (turning, grinding, drilling, milling) [2, 7, 9, 10].

In principle, the methods of analysis and synthesis of contemporary automatic control systems are based on solving a problem with notable simplification of the physical and mathematical relations that characterise the processes included in the system. This is largely a result of imperfection of the research apparatus employed, and of complexity of acquired in prior information on the dynamic and static characteristics of the object of control and of external interference factors. The existing methods of synthesis of automatic control systems take into account the possible uncertainty concerning the character and extent of interference only to a limited degree. Even less developed are the issues of scatter of control object parameters, though at present considerable attention is devoted to them.

Designing of ACS capable of operating at uncontrolled variability of control object parameters has led to the appearance adaptive control (AC) in machine tools [1, 3, 8]. In spite of its flexibility and possibility of assurance of required quality of transition processes for a broad range of objects, in many cases the application of AC on machine tools is difficult due to the necessity of continuous measurement of characteristics of technological systems and interference.

Reserves for increasing the accuracy and quality of machining may be defined in the design of optimum structures of control systems, as the ACS of elastic deformations (by means of relative moves of the cutting tool and the machined part) of the DS is static in terms of both control effects and interference, and their change causes errors of mutual positioning of the part and the cutting edge [4, 8].

The theoretical foundations for improvement of machining quality with adaptive control, considered in this report, are based on mathematical description of the object of control in the DS presented in references [2, 4, 5, 7, 10] and may be applied in DS analyses described mathematically differently than the presented herein and taking into account a specific character of the control object.

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

2. Some issues of invariability in the control of elastic deformations of the dynamic system

In principle, systems of automatic control of technological processes comprise the object of control - the dynamic system of the machine tool, made up of the MDE system of the machine tool and the working processes involved in their interaction, and the regulator. Analysis of connections within the dynamic system of the machine tool indicates that it is a multi-contour closed system, and the interaction of the fundamental components is directed by the interactions between the working processes, taking place via the MDE system only MDS [2, 6]. Separation of the zones where the working processes take place, by means of the elements of the elastic system permits relative simple transition to equivalent concepts of the dynamic system of the machine tool as a single-contour system.

The transfer function of an equivalent MDE system has the form:

$$G_{us}(s)=\frac{y(s)}{F_p(s)}=\frac{K_s}{T_1^2 s^2+T_2 s+1}$$

where: $y(s)$, $F_p(s)$ – presentation, acc. to Laplace, of elastic deformations and force of machining;

$K_s=1/C_s$, $T_1=\sqrt{m_y/C_s}$, $T_2=n_y/C_s$ – coefficient of proportionality and time constants of MDE system of ma chine tool;

m_y, n_y, C_y – reduced mass, coefficient of attenuation and rigidity of MDE system.

The transfer functions of elements of adaptive control system (Fig. 1a) have the following form:

- for the MDE system, relative to the control effect

$$G_s^s(s)=\frac{y(s)}{F_p(s)}=\frac{K_s}{T_{1s}s^2+T_{2s}s+1}$$

- for the MDE system, relative to interference effect

$$G_s^s(s)=\frac{y(s)}{F_p(s)}=\frac{K_s}{T_{1s}s^2+T_{2s}s+1}$$

- for the process of machining

$$G_{sk}(s)=\frac{F_p(s)}{\mathrm{d}(s)}=\frac{K_{sk}}{T_{sk}s+1},$$

- for the executive mechanism

$$G_w(s)=\frac{K_w(s)}{\mathrm{e}(s)}=\frac{K_w}{T_w s+1},$$

where: $T_{1s}=m_y/C_s$, $T_{2s}=n_y/C_s$, $K_s=1/C_s$, $K_f=K_{1s}/C$, T_w, K_w – time constant and coefficient of proportionality of transfer function of the executive mechanism.

The accuracy of operation of the system of adaptive control of elastic deformations of DS is characterised by an error e_s equal to [2, 7]:

$$\mathrm{e}_s=\mathrm{e}_s^o+\mathrm{e}_f, \qquad (1)$$

where: e_s^0 – error caused by control effect $y^0(t)$,

e_f – error caused by interference effect $f(t)$.

Errors e_s^o and e_f are determined from the relation:

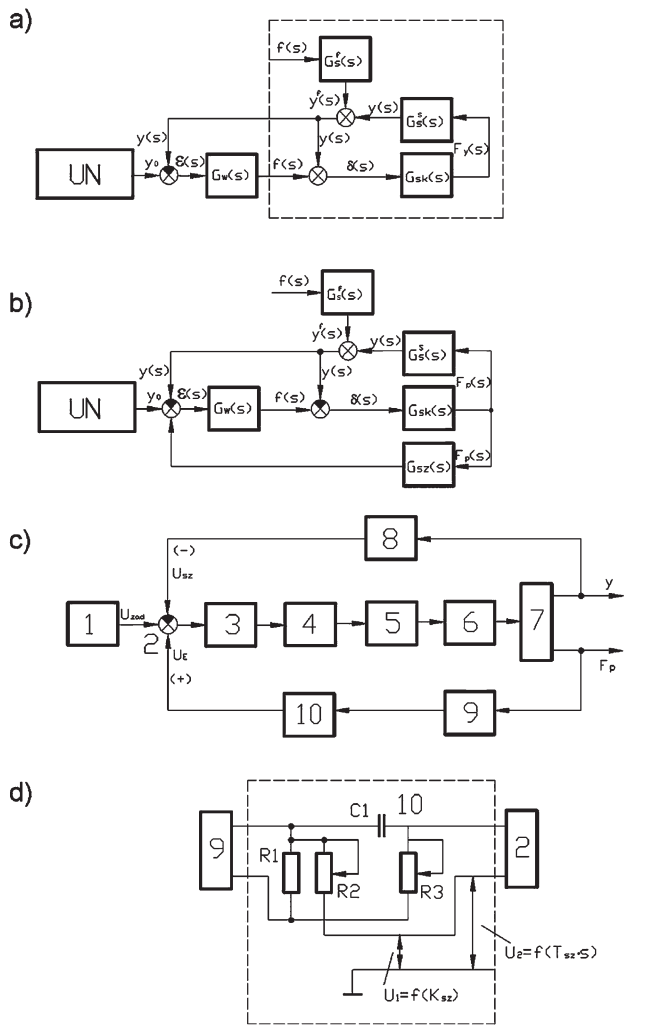


Fig. 1. Block diagram of the system of stabilisation of elastic deformations of technological system – a; schematic of system of adaptive control of elastic deformations of technological system – b; functional schematic of adaptive control – c; node of positive feedback relative to the force of machining – d

$$\mathrm{e}_s{}^0(s)=\frac{1+G_{sk}(s)\cdot G_s^s(s)}{1+G_s^s(s)\cdot G_{sk}(s)+G_s^s(s)\cdot G_{sk}(s)\cdot G_w(s)}y^0(s)=\frac{\sum_{i=0}^{4}b_i s^i}{\sum_{j=0}^{4}a_j s^j}y^0(s), \qquad (2)$$

$$\mathrm{e}_f(s)=\frac{G_s^f(s)}{1+G_s^s(s)\cdot G_{sk}(s)+G_s^s(s)\cdot G_{sk}(s)\cdot G_w(s)}f(s)=\frac{\sum_{i=0}^{2}d_i s^i}{\sum_{j=0}^{4}a_j s^j}f(s), \qquad (3)$$

where: $a_0=K_sK_{sk}+K_sK_{sk}\cdot K_w$; $b_0=1+K_{sk}K_s$;

$a_1=b_1=T_w+K_sK_{sk}K_w+T_{sk}+T_{2s}$;

$a_2=b_2=T_{sk}T_w+T_{2s}T_w+T_{2s}T_{sk}+T_{1s}$;

$a_3=b_3=T_{1s}T_w+T_{2s}T_{sk}T_w+T_{1s}T_{sk}$;

$a_4=b_4=T_{1s}T_{sk}T_w$;

$$d_0 = K_s;$$
$$d_1 = K_s(T_{sk} + T_w);$$
$$d_2 = K_s T_{sk} T_w.$$

The determined values of the errors (2) and (3) can be presented as [8]:

$$e_{y^0}(t) = C_0 y^0(t) + C_1 \frac{dy^0(t)}{dt} + \frac{C_2}{2!} \cdot \frac{d^2 y^0(t)}{dt^2} + \frac{C_3}{3!} \cdot \frac{d^3 y^0(t)}{dt^3} + \frac{C_4}{4!} \cdot \frac{d^4 y_0(t)}{dt^4},$$

$$e_f(t) = C_0' f(t) + C_1' \frac{df(t)}{dt} + \frac{C_2'}{2!} \cdot \frac{d^2 f(t)}{dt^2} + \frac{C_3'}{3!} \cdot \frac{d^3 f(t)}{dt^3} + \frac{C_4'}{4!} \cdot \frac{d^4 f(t)}{dt^4}, \quad (4)$$

where C_i, C_i' (i = 0, 1, 2, ..., 4) – coefficients of error, characterising the accuracy of operation of the control system and dependent on its structure.

In the structure of the control system (Fig. 1), the values of the coefficients of error are defined by the following relations:

$$C_0 = \frac{b_0}{a_0}, \quad C_1 = \frac{b_1 - a_1 C_0}{a_0}, \quad \frac{C_2}{2} = \frac{b_2 - a_1 C_1 - a_2 C_0}{a_0},$$

$$C_0' = \frac{d_0}{a_0}, \quad C_1' = \frac{d_1 - a_1 C_0'}{a_0}, \quad \frac{C_2'}{2} = \frac{d_2 - a_1 C_1' - a_2 C_0'}{a_0}.$$

It should be noted that the main error is constituted by the error coefficients C_0, C_0' and C_1, C_1', referred to as the static and speed errors. For a 16K20 machine tool, with system parameters of:

$$T_{1s} = 1{,}6 \cdot 10^{-6} s^2, \; T_{2s} = 1{,}2 \cdot 10^{-4} s, \; K_s = 1{,}6 \cdot 10^{-6} mm/N, \quad (5)$$

$$T_{sk} = 10^{-3} s, \; T_w = 10^{-3} s, \; K_f = 2 \cdot 10^{-4} mm/N, \; K_{sk} = 1$$

based on data (5), the values of the error coefficients were determined analytically, as follows:

$$C_0 = 0{,}32 mm^{-1}, \; C_1 = 1{,}1 \cdot 10^{-3} s/mm, \; C_2 = 0{,}62 \cdot 10^{-6} s^2/mm,$$
$$C_0' = 1{,}5 \cdot 10^{-3} mm^{-1}; \; C_1' = -1{,}1 \cdot 10^{-6} s/mm, \; C_2' = -2{,}73 \cdot 10^{-9} s^2/mm.$$

Total or partial elimination of those errors would permit improvement of the accuracy of operation of the control system and, correspondingly, of the machining of the part.

The problem posed can be solved through the introduction of an additional positive feedback relative to the force of machining, with transmittance $G_{sz}(s)$, into the system of control of elastic deformations of the DS (Fig. 1b) [2].

In this case, error $e_s^0(s)$ is determined from the relation:

$$e_s^0(s) = \frac{1 - G_{sk}(s) G_{sz}(s) G_w(s) + G_{sk}(s) G_s^s(s)}{1 - G_{sk}(s) G_{sz}(s) G_w(s) + G_s^s(s) G_{sk}(s) G_w(s)} \; y^0(s). \quad (6)$$

As follows from expression (6), the error introduced in the system of control of elastic deformations of the TS by the effect $y^O(t)$ can be totally eliminated if the structure and parameters of the transfer function are presented as follows:

$$G_{sz}(s) = \frac{1 + G_{sk}(s) G_s^s(s)}{G_{sk}(s) G_w(s)}. \quad (7)$$

Expression (7) can be called the condition of total invariability of the control system relative to the control effect $y^O(t)$. If this condition is met, all the error coefficients C_i (i = 1, 2, 3, 4) are equal zero.

Taking into account that meeting the condition (7) basically leads to systems that are physically unrealisable, and that sufficient accuracy can be achieved in practice when only the static error coefficients C_0 or C_0 and C_i are equal to zero, it is enough to meet the conditions of $b_0 = 0$ and $b_1 = 0$.

However, with $b_1 = 0$ also coefficient a_1 equals zero, which causes that the system loses stability. Therefore, in the case under consideration, it is only possible to eliminate the static error C_0.

Error C_0 equal zero can be achieved with means that are technologically simple enough, through the introduction of an additional positive feedback relative to the force of machining, with transfer function of $G_{sz}(s) = K_{sz}$.

The value of the coefficient of proportionality of the feedback is determined from relation (7)

$$b_0 = 1 + K_{sk} K_s - K_{sk} K_{sz} = 0, \quad (8)$$

that is $K_{sz} = (1 + K_{sk} K_s) / K_{sk}$,

however – to maintain stability – it must be lower than 1.

If condition (8) is met, in the control system with parameters (5) considered earlier, the coefficients of error are equal to

$$C_0 = 0; \; C_1 = 2{,}33 \cdot 10^{-2} s/mm.$$

Error introduced by an interference effect can be presented as:

$$e_f(s) = \frac{G_s^f(s)\left[1 - G_{sk}(s) G_w(s) G_{sz}(s)\right]}{1 - G_{sk}(s) G_w(s) G_{sz}(s) + G_s^s(s) G_{sk}(s) G_w(s) + G_s^s(s) G_{sk}(s)}. \quad (9)$$

Relation (9) can be also used to obtain the conditions of total invariability of the system under the effect of an interference factor

$$G_{sz}(s) = \frac{1}{G_{sk}(s) G_w(s)}. \quad (10)$$

By analogy, condition (10) will be met if the structure of the transfer function of the feedback is presented as:

$$G_{sz}(s) = K_{sz} + T_{sz} \cdot s,$$

and its parameters will be selected from the equations:

$$d_0 = 1 - K_{sz} K_{sk} K_w = 0,$$

$$d_1 = T_{sk} + T_w - T_{sz} K_{sz} K_w = 0,$$

hence $K_{sz} = \dfrac{1}{K_{sk} K_w}, \quad T_{sz} = \dfrac{T_{sk} - T_w}{K_{sk} K_w}.$

In this case, the coefficients of error in the example under consideration will be equal to:

$$C_0 = 0, \; C_1 = 0, \; C_2 = 6{,}25 \cdot 10^{-9} s^2/mm.$$

Fig. 1c presents a block diagram of a system of adaptive control realizing the method of adjustment of settings of the TS [8]. The control system incorporates, in series connection, the elastic deformation

regulator 1, comparing element 2 to one of the inputs, connected to which is the elastic deformation sensor 8, correcting element 3, amplifier 4, power transducer 5, longitudinal feed motor 6, technological system 7, machining force sensor 9, node of positive feedback 10. Node 10, of the positive feedback relative to the machining force F_p (Fig. 1d) is made as a differential chain of condenser C_1, resistor a R3 and a voltage divider - resistors R1 and R2 .

In the process of work, preliminary determination is made of the initial value of elastic deformation y, in the form of signal U_z, by means of regulator 1. The true value of elastic deformation y is measured by sensor 8, and the result of the measurement, in the form of voltage U_{sz}, is summed up algebraically on the comparing element 2 with the regulated voltage U_z. At the same time, sensor 9 measures the change in the machining force F_p, caused by changes in machining conditions (material hardness of semi-finished product, allowance amount, width of machined layer, initial error). The change in the value of machining force F_p in the technological system causes, in turn, deviation of the elastic deformation y from the adopted value.

Signal from sensor 9, of machining force, is supplied to the input of node 10, of the positive feedback, from which signals $U_1 = f(K_{sz})$ are taken off from the output of the voltage splitter, on resistors R1 and R2, and $U_2 = f(T_{sz} \cdot s)$ from the output of the differential chain C1 and R3.

The setting of parameters K_{sz} and T_{sz} and the corresponding signals U_1 and U_2 for a specific technological system is realized by means of the adjustable resistors R2 and R3. Signals $U_1 = f(K_{sz})$ and $U_2 = f(T_{sz} \cdot s)$ are supplied to the inputs of the comparing element 2, where they are algebraically summed up with U_z and U_{sz}. The error signal $U_e = U_z + U_1 + U_2 - U_{sz}$, via the correction element 3, is supplied to the input of amplifier 4, and then, through the power transducer, to the longitudinal feed motor 5. Change in the rotation speed of motor 6 causes a change in the value of the longitudinal feed that is a control effect for the technological system 7, and thus a correction is made to the relative positioning of the machined part and the cutting edge, taking into account the change in elastic deformation y and the machining force F_p .

3. Control of parameters of elastic-deformable state of low-rigidity shafts in turning

The above considerations concerning the creation of a system of adaptive control of accuracy parameters in machining with control effects in the form of longitudinal feed can also be generalised onto the dynamic systems of profiling of elastic-deformable low-rigidity shafts.

References [8, 9] present a mathematical description of systems in turning and grinding of elastic-deformable low-rigidity shafts, the dynamic properties of linearized models being approximated by means of transfer functions of typical dynamic elements. The resultant models of DS and of the parameters of the object of control permit substantiated realization of the search for optimum algorithms of control, selection of control system structure, and synthesis of corrective devices.

Fig. 2 presents a generalised structural schematic of the object of control for the case of controlling the elastic deformable state of low-rigidity parts through the application of tensile force F_{x1}. The transfer function of the dynamic system, taking into account the assumptions and results of theoretical and experimental studies, can be presented as:

$$G_4(s) = \frac{1 + K_k K_y m_y \left(1 - e^{-st}\right)}{1 + K_{bz} K_z n_z + K_{xy} n_x + K_{yy} n_y + \left(1 - e^{-st}\right)} \times \frac{1}{\left[k_k k_y m_y \left(1 + K_{xy} n_x + K_{bz} K_z n_z + K_{yy} n_y\right) - \left(K_k - K_x n_x\right)\left(K_{bz} K_z n_z + K_{yy} m_y\right)\right]}. \quad (11)$$

Fig. 3 presents a structural schematic of a system of adaptive control of elastic deformations of an elastic-deformable part in DS. with an additional feedback $G_{sz}(s)$ relative to the machining force F_p, that realizes the condition (7) of invariability in order to eliminate the static error related to the control effect $y_0(s)$ with positive feedback and to increase stability, speed of operation and insensitivity to a change in material allowance with negative feedback. In Fig. 3 the expression for $G_5(s)$ can be presented as:

$$G_5(s) = \frac{\left(1 - e^{-st}\right)\left[m_y\left(K_x n_x - K_k\right) - K_k K_y m_y n_y\right] - n_y}{1 + K_k K_y m_y \left(1 - e^{-st}\right)}. \quad (12)$$

Taking into account relations (11) and (12), the transfer function of the corrected control system can be written as:

$$\Phi_{sk}(s) = \frac{1 + K_{bz} K_z n_z + K_{xy} n_x + K_{yy} n_y + n_y K_{F_{x1}} G_w(s) + \left(1 - e^{-st}\right)\left[K_k K_y m_y \times \left(1 + K_{xy} n_x + K_{bz} K_z n_z + K_{yy} n_y + n_y K_{F_{x1}} G_w(s) G_{sz}(s)\right) + \left(1 - e^{-st}\right) \times \left[K_k K_y m_y \left(1 + K_{xy} n_x + K_{bz} K_z n_z + K_{yy} n_y + K_{F_{x1}} G_w(s)\right)\left(1 + n_y G_{sz}(s) + K_{xy} n_x\right) + \left(K_x n_x - K_k\right)\left(K_{bz} K_z n_z + m_y K_{yy} - m_y K_{F_{x1}} G_w(s) G_{sz}(s)\right)\right]}{1 + K_{bz} K_z n_z + K_{xy} n_x + K_{yy} n_y + K_{F_{x1}} G_w(s)\left(1 + n_y G_{sz}(s)\right) + \left(1 - e^{-st}\right) \times \left[K_k K_y m_y \left(1 + K_{xy} n_x + K_{bz} K_z n_z + K_{yy} n_y + K_{F_{x1}} G_w(s)\right)\left(1 + n_y G_{sz}(s) + K_{xy} n_x\right) + \left(K_x n_x - K_k\right)\left(K_{bz} K_z n_z + m_y K_{yy} - m_y K_{F_{x1}} G_w(s) G_{sz}(s)\right)\right]}, \quad (13)$$

and the static error of the system relative to the control effect is determined from the relation:

$$e_y(s) = y_0(s) \frac{1 + K_{bz} K_z n_z + K_{xy} n_x + K_{yy} n_y - n_y K_{F_{x1}} G_w(s) G_{sz}(s)}{1 + K_{bz} K_z n_x + K_{yy} n_y + K_{yy} n_y + K_{F_N}(s) G_w(s)\left(1 + n_y G_{sz}(s)\right)}. \quad (14)$$

As follows from relation (14), the static error $e_y(s)$, introduced into the system of control of parameters of the elastic-deformable state of parts in DS. by the controlling effects, can be eliminated if the structure and the parameters of transfer function of positive feedback $G_{sz}(s)$ is selected as follows:

$$G_{sz}(s) = \frac{1 + K_{bz} K_z n_z + K_{xy} n_y + K_{yy} n_y}{n_y K_{F_{x1}} G_w(s)}. \quad (15)$$

If we leave out the effect of component F_f of machining force on the increase of elastic deformations along coordinate y, the structural schematic of the object of control – dynamic system of elastic-deformable low-rigidity shafts can be transformed to the form presented in Fig. 4a, and the transfer function of the object is then defined by the relation:

$$G_6(s) = \frac{1 + m_y K_y K_k \left(1 - e^{-st}\right) \times}{1 + K_{bz} K_z n_z + K_{yy} n_y + \left(1 - e^{-st}\right) \times} \frac{\times 1}{\times \left[m_y K_y K_k \left(1 + K_{bz} K_z n_z + K_{yy} n_y\right) + \left(K_{bz} K_z m_z + K_{yy} m_y\right)\left(K_k - n_x K_x\right)\right]}. \quad (16)$$

The structural schematic of AC with introduced feedback relative to the force of machining is presented in Fig. 4b, where the transfer function $G_5(s)$ is defined by relation (12).

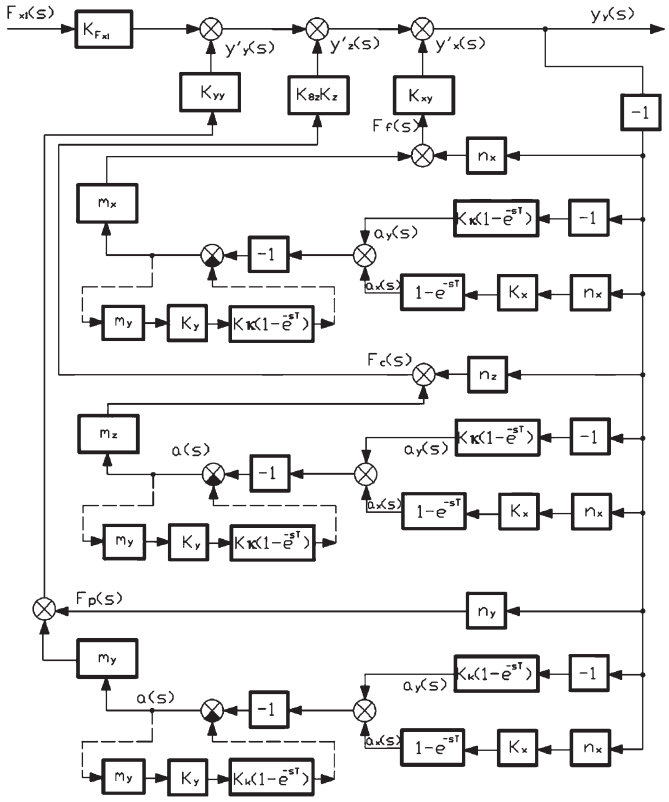


Fig. 2. Structural schematic of dynamic system in turning of elastic-deformable low-rigidity shafts

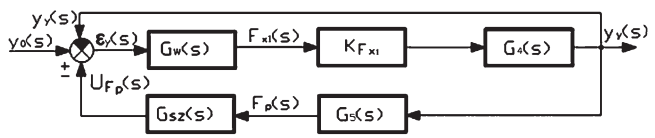


Fig.3. Structural schematic of adaptive control system with additional positive feedback relative to the machining force

The static error of the control system, introduced by a controlling effect, is defined by the relation:

$$e_{y1}(s)=y_0(s)\frac{1+K_{bz}K_z n_z+K_{yy}n_y-n_yK_{F_{x1}}G_w(s)G_{sz1}(s)}{1+K_{bz}K_z n_z+K_{yy}n_y+K_{F_{x1}}G_w(s)\left[1+n_yG_{sz1}(s)\right]}. \quad (17)$$

To eliminate the static error $e_{y1}(s)$ introduced into the system by the controlling effects, the structure and parameters of transfer function of positive feedback $G_{sz1}(s)$ should be defined as follows:

$$G_{sz1}(s)=\frac{1+K_{bz}K_z n_z+K_{yy}n_y}{n_yK_{F_{x1}}G_w(s)}. \quad (18)$$

In the case when we leave out the effect of elastic deformations along coordinates z and x on the change in machining depth (along coordinate y), the structural schematic for a specific model of the technological system of turning of elastic-deformable shaft can be presented as in Fig. 5a, where the transfer function $G_5(s)$ is defined by the relation (12). Fig. 5b presents the structural schematic of cor-

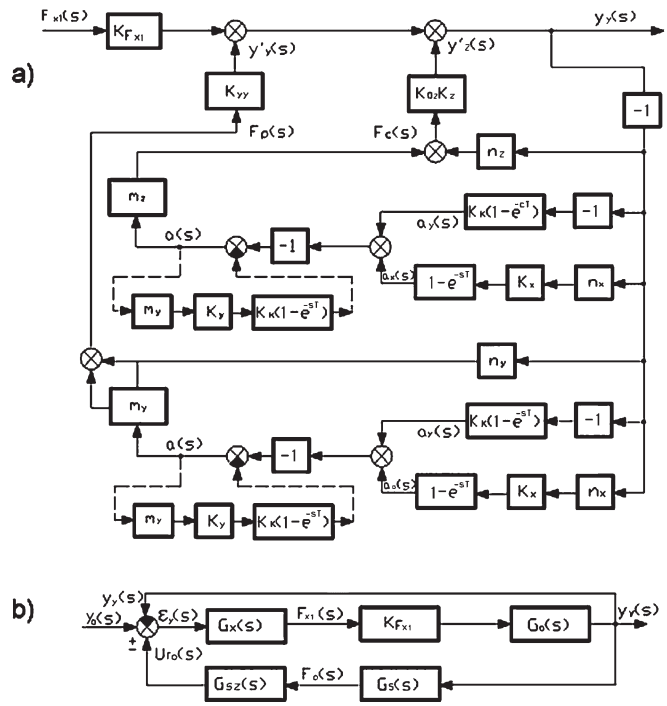


Fig.4. Structural schematic of control object without the inclusion of the effect of component F_p of machining force on increase of elastic deformation along coordinate y – a; structural schematic of adaptive control system with feedback relative to machining force F_p - b

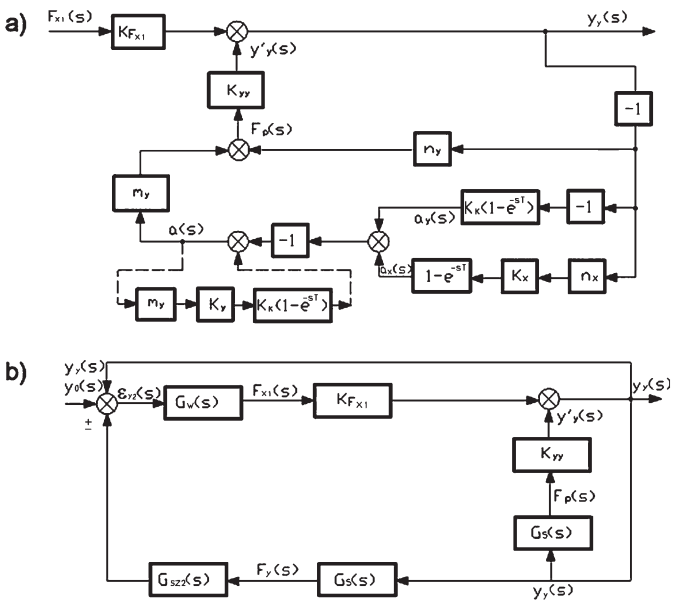


Fig. 5. Structural schematic of specific model of technological system of turning of elastic-deformable low-rigidity shafts – a; structural schematic of adaptive control system – b

rected system of adaptive control of the parameters of the elastic-deformable state of a low-rigidity shaft, that permits the elimination of the static error with the introduction of an additional positive feedback. The transfer function for the corrected system for the specific model is defined by the relation:

$$\Phi_{sk}(s)=\frac{1+K_{yy}n_y+n_yK_{F_{x1}}G_w(s)G_{sz2}(s)+m_y\left(1-e^{-st}\right)}{1+K_{yy}n_y+K_{F_{x1}}G_w(s)\left[1+n_yG_{sz2}(s)\right]+m_y\left(1-e^{-st}\right)}\times$$
$$\times\frac{\left[K_kK_y+\left(K_{yy}+K_{F_{x1}}G_w(s)G_{sz2}(s)\right)\left(n_yK_{yy}K_k-K_xn_x+K_k\right)\right]}{\left[K_kK_y\left(1+K_{F_{x1}}G_w(s)\right)+\left(K_{yy}+K_{F_{x1}}G_w(s)G_{sz2}(s)\right)\left(K_kK_{yy}n_y-K_xn_x+K_k\right)\right]}. \quad (19)$$

The error of the control system introduced by a controlling effect is defined from the relation:

$$e_{y2}(s)=y_0(s)\frac{1+K_{yy}n_y-n_yK_{F_{x1}}G_w(s)G_{sz2}(s)}{1+K_{yy}n_y+K_{F_{x1}}G_w(s)\left[1+n_yG_{sz2}(s)\right]}. \quad (20)$$

In the case of determining the structure and parameters of feedback relative to the force of machining in accordance with the relation:

$$G_{sz2}(s)=\frac{1+K_{yy}n_y}{n_yK_{F_{x1}}G_w(s)}, \quad (21)$$

the static error $e_{y2}(s)$ relative to the controlling effect assumes the value of zero.

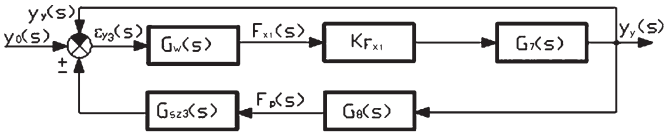


Fig. 6. Structural schematic of system of adaptive control of parameters of elastic-deformable state in longitudinal grinding

4. Control of parameters of elastic-deformable state of low-rigidity shafts in grinding

The generalised model of dynamic system of turning of elastic-deformable parts, taking into account the specific character of profiling the shear section in grinding of elastic-deformable low-rigidity shafts with longitudinal feed, was used to derive the structure of the object of control and the mathematical description of the system [9].

The transfer function of DS in longitudinal grinding is defined by the relation:

$$G_7(s)=\frac{1+m_xK_x\left(1-e^{-st}\right)}{1+K_{xy}n_x+K_{yy}n_y+\left(1-e^{-st}\right)\left[m_xK_x-K_{yy}K_x\left(m_yn_x-m_xn_y\right)\right]}. \quad (22)$$

Fig. 6 presents the structural schematic of a system of adaptive control of the parameters of elastic-deformable state of low rigidity shaft in TS in longitudinal grinding, where:

$$G_8(s)=-\frac{n_y+\left(1-e^{-st}\right)\left(m_yn_xK_x-m_xn_yK_x\right)}{1+m_xK_x\left(1-e^{-st}\right)}. \quad (23)$$

The transfer function of the corrected control system, taking into account relations (22) and (23), assumes the form of:

$$\Phi_{sk}(s)=\frac{1+K_{xy}n_x+K_{yy}n_y+n_yK_{F_{x1}}G_w(s)G_{sz3}(s)+\left(1-e^{-st}\right)\times}{1+K_{xy}n_x+K_{yy}n_y+K_{F_{x1}}G_w(s)\left[1+n_yG_{sz3}(s)\right]+\left(1-e^{-st}\right)\times}$$
$$\frac{\times\left[m_xK_x+\left(K_{yy}-G_w(s)G_{sz3}(s)K_{F_{x1}}\right)\left(n_y-K_xn_xm_y\right)\right]}{\times\left[m_xK_x\left(1+G_w(s)K_{F_{x1}}\right)+\left(K_{yy}-G_w(s)G_{sz3}(s)K_{F_{x1}}\right)\left(n_y-K_xn_xm_y\right)\right]} \quad (24)$$

If the structure and parameters of the transfer function of positive feedback $G_{sz3}(s)$ are determined from the relation:

$$G_{sz3}(s)=\frac{1+K_{xy}n_x+K_{yy}n_y}{n_yK_{F_{x1}}G_w(s)}, \quad (25)$$

then, as follows from relation (20), the static error introduced in the adaptive control system in longitudinal grinding by the controlling effect, $y_0(s)$c, can be eliminated.

5. Conclusion

Analysis of functioning of systems of automatic control of elastic deformations in the technological system shows that under dynamic conditions they are static relative to the effect of both controlling and interference effects. This leads to errors of the relative positioning of the machined part and the cutting edge. Optimum structures of control systems, including the developed systems of adaptive control, permit improvement of the quality (accuracy) of machining.

The authors have developed a method for the correction of the setting of the technological system through the application of adaptive control of longitudinal feed and additional force effects, inducing the elastic-deformable state through the introduction into the control system of an additional positive feedback relative to the machining force, permitting the elimination of static errors of controlling and interference effects in the control of parameters of accuracy and quality.

The conditions were developed for the determination of the structure of parameters of the additional positive feedback relative to the machining force, that guarantee the elimination of static errors and impart to the structure adaptive properties.

References

1. Halas W, Taranenko V, Swic A, Taranenko G. Investigation of influence of grinding regimes on surface tension state. Berlin, Heidelberg: Springer – Verlag. Lecture Notes In Artificial Intelligence, 2008; 5027: 749–756.
2. Marchelek K.: Dynamika obrabiarek, WNT: Warszawa, 1991.
3. Kujan K. Badania i analiza powtarzalności rozkładu odchyłek geometrycznych w procesie obróbki skrawaniem. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2008; 3(39): 45–52.
4. Ratchev S, Liu S, Huang W, Becker A A. A flexible force model for end milling of low-rigidity parts. Journal of Materials Processing Technology. Proceedings of the International Conference in Advances in Materials and Processing Technologies. 2004; 153-154: 134-138.
5. Ratchev S, Liu S, Huang W, Becker A A. Milling error prediction and compensation in machining of low-rigidity parts. International Journal of Machine Tools and Manufacture 2004; 15(44): 1629-1641.

6. Ratchev S, Govender E, Nikov S. Towards deflection prediction and compensation in machining of low-rigidity parts. Proceedings of the Institution of Mechanical Engineers - Part B 2002; 1(216): 129-134.
7. Świć A, Taranenko W, Szabelski J. Modelling dynamic systems of low-rigid shaft grinding. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2011, 2 (50): 13 – 24.
8. Taranenko W, Świć A. Urządzenia sterujące dokładnością obróbki maszyn o małej sztywności. Lublin: Politechnika Lubelska, 2006.
9. Taranenko W, Taranenko G, Szabelski J, Świć A. Identyfikacja układu dynamicznego szlifowania wałów o małej sztywności. Modelowanie Inżynierskie 2008, 4(35): 115 – 130.
10. Taranenko G, Taranenko W, Świć A, Szabelski J. Modelling of dynamic system of low-rigidity shaft machining. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2010, 4 (48): 4–15.

Dr hab. inż. Antoni ŚWIĆ, prof. nadzw.
Institute of Technological Information Systems
Lubliln University of Technology
ul. Nadbystrzycka 36, 20-618 Lublin, Poland
e-mail: a.swic@pollub.pl

Prof. dr hab. inż. Wiktor TARANENKO
Institute of Technological Information Systems
Politechnika Lubelska
ul. Nadbystrzycka 36, 20-618 Lublin, Poland
e-mail: wtaran_ark@mail.ru

Article citation info:
TOMPOROWSKI A, OPIELAK M. Structural features versus multi-hole grinding efficiency. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 222-227.

Andrzej TOMPOROWSKI
Marek OPIELAK

STRUCTURAL FEATURES VERSUS MULTI-HOLE GRINDING EFFICIENCY

KONSTRUKCYJNE KSZTAŁTOWANIE WYDAJNOŚCI ROZDRABNIANIA WIELOOTWOROWEGO*

The purpose of this paper has been to describe the influence of design features and properties of disintegrated grain biomaterials on the dynamics of the process efficiency. It has been determined that the basis for improved functionality of the grinding machine is the analysis of the influence of the effective component of gravitational force and angle of repose of the grain biomass on the dynamics of changes in the efficiency of the five-disc, multi-hole RWT-5KZ grinder. Reasonable efficiency may be obtained by means of purposeful control of cross-sections and volumes of grinding holes, i.e. design features of multi-hole multi-disc unit. The actions will however bring the planned benefits only when a mathematical description is developed for the flow of the disintegrated grains (biomass grains) through the working space of the multi-disc grinder as a resultant variable of the design and operation of the working unit. The search for design solutions of the units that grind grain, leading to efficient processing justify the research into the improvement of the theory and design of grinding machines.

***Keywords**: grinding, biomaterials, efficiency.*

W pracy podjęto próbę opisu wpływu cech konstrukcyjnych oraz właściwości rozdrabnianych biomateriałów ziarnistych na dynamikę wydajności procesu. Uznano, że podstawą do poprawy funkcjonalności działania maszyny rozdrabniającej jest przeprowadzenie rozpoznania wpływu efektywnej składowej siły grawitacji i kąta usypu biomasy ziarnistej, na dynamikę zmian wydajności pięciotarczowego, wielootworowego rozdrabniacza RWT-5KZ. Racjonalna wydajność może być osiągnięta, między innymi, na drodze celowego sterowania przekrojami i objętościami otworów uczestniczących w rozdrabnianiu, czyli cechami konstrukcyjnymi wielotarczowego zespołu wielootworowego. Aby jednak działania te przyniosły planowane korzyści, konieczne staje się opracowanie opisu matematycznego przepływu rozdrabnianych ziaren zbóż (ziaren biomasy) przez przestrzeń roboczą rozdrabniacza wielotarczowego, jako zmiennej wynikowej konstrukcji i działania zespołu roboczego. Poszukiwania rozwiązań konstrukcyjnych zespołów rozdrabniających ziarna zbóż, prowadzące do wydajnych procesów przetwórczych, uzasadniają podjęcie badań nad doskonaleniem teorii i konstrukcji rozdrabniaczy.

***Słowa kluczowe**: rozdrabnianie, biomateriały, wydajność.*

1. Introduction

Innovations in, designing and testing of drive systems of grinders constitute a big challenge for engineers. Experience gained over a few decades shows that the creation of an integrated, multi-purpose system is not simple [1, 3, 4, 5]. Grinding has become a common process popular in industry and everyday life.

Functional model of grain material grinders must take into account the impact of design features of working elements on the operation and the basis for the physical interactions in material processing, according to the principles of mechanics. The study focused on the relationship between design features of the multi-disc motion/process unit and performance characteristics of the machine [3, 8].

Model unevenness, as the exemplification of performance characteristics of multi-edge grinders for biomass grains depends on power, torque (useful work) and the angular velocity (also angle of rotation, acceleration of the cutting edges of the process unit in time) and movement (transport) of the feed material in working space of the grinding machine.

As regards multi-edge grinding of grain biomass, design features of the discs (Fig. 1) must ensure performance of two basic motion functions: movement and comminution in the inter-hole space. The results of previously conducted experiments confirmed the validity of the theoretical assumptions made and correctness of mathematical relations developed [9, 10].

According to the literature, research was carried out into e.g. the effect of a rotational mechanical medium on the behaviour of pieced materials inside that medium. In 1994 Walton made an attempt to determine the so-called dynamic angle of repose, taking into account the values of the friction factor during the cutting of granulate and, as a consequence, its deflection [12]. In 2007 Ulrich S., Schroter M. and Swinney H. investigated the influence of the friction factor on the behaviour of the mixture comprising spherical particles of materials [11]. As yet, the literature is exclusive of any researches into the influence of the gravitational force effective component and the angle of repose of biomass grains on the dynamics of changes in the efficiency of multi-disc grinding.

The primary objective of this study was to describe phenomena, process variables, relations of the multi-disc grinding, and especially to answer the question: do the effective gravitational force and angle of repose influence the dynamics and efficiency of five-disc, multi-hole 5KZ RWT grinder for grain biomass?

2. Geometric form of biomass grains

Biomass grains with stabilized moisture parameters and grain dimensions were used as the feed material in the analysis of volumetric efficiency. An assumption was

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

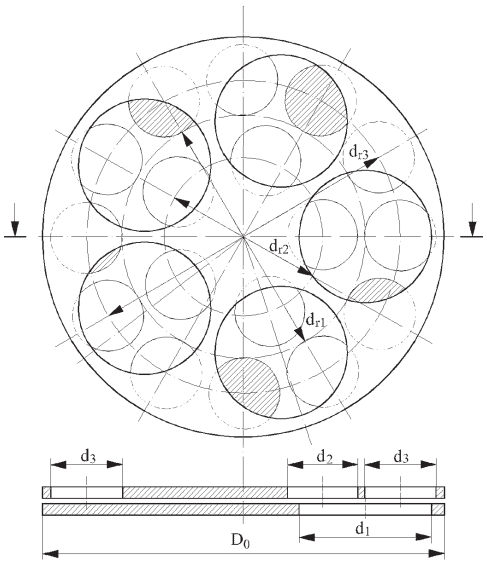


Fig. 1. *Design features, cross-sectional area and effective grinding surface for two grinding working discs between cutting edges [5]: D_0–outer disc diameter, d_{r1}– diameter of working holes mounting in the preceding disc, d_{r2}– mounting diameter of working holes inner row in the next disc, d_{r3}– mounting diameter of working holes outer row in the next disc, d_1- diameter of working hole in the preceding disc, d_2, d_3–diameters of working holes in the next disc.*

made, therefore, that for the purpose of this analysis an output model of the feed material is the material selected by dimensions with a repeatable constant size of an individual grain. Taking into consideration shapes of the basic grains (Fig.2), an assumption was made that biomass grains take the position in the holes of the working discs along their longer axis, perpendicularly to the inter-disc cutting plane.

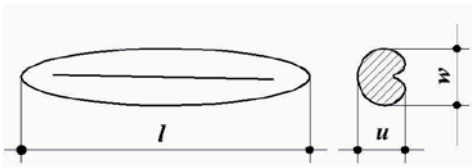


Fig. 2. Model of the geometric form of the basic grains: l, u, w – maximum dimensions of grain within three planes

Arrangement of the disintegrated biomass within the working space of the multi-hole grinder has been described with a probability distribution for grain length. Because in the holes of the same disc the material is described with the same grain-size state and is subject to the same cutting process in each hole, its state for the purpose of this analysis will be indexed with the cut number (m) and disc number (n):

$$\rho_n^m : [0, l_{\max}] \to [0,1], \quad \int_0^{l_{\max}} \rho_n^m dl = 1. \tag{1}$$

Since the desirable consequence of technological movement of the disintegrated material is its transfer to the next disc, the process results in the quasi-cutting at the disc edges. The initial state of the material in the first layer (disc) before the first quasi-cut is ρ_0^0 and constitutes an input state of the feed material in the working space of the analysed grinder. This state is hereinafter described by a function focused on value l_{max}. Efficiency-related state of the material during grinding changes as a result of two factors, mechanisms: grinding (quasi-cutting) and the removal of grains of the desired size from the inter-disc space of the grinding device operating unit.

3. Grinding efficiency model

When holes of two adjacent discs meet (Fig.1 and Fig.3) and their common cross-sectional area begins to increase (0, max), as a result of the movement of the feed material, the next grinding hole (in the next disc) is filled with some of the material from the preceding hole (in the preceding disc). In order for the cutting process to be efficient, the following conditions must be met:

- the hole of the preceding disc must be filled completely when it meets the next hole (when the common cross-sectional area of the discs appears),
- the hole of the next disc must be completely filled when the common cross-sectional area of the holes begins to decrease (the maximum common temporary cross-sectional area of the two grinding holes).

The two conditions must be met:

1. The volume of the feed material in the preceding hole must always be smaller than the space available in the next hole. Otherwise, undesirable idle movement of the feed material may occur, i.e. the material may move from the hole of the preceding disc to the hole of the next disc without being cut (quasi-cut) along the whole length of the device. Because some of the material as small particles is removed through the inter-disc space (inter-disc gap is a required design feature of this type of grinder), the volume of holes in the next disc should always be smaller than in the preceding disc. This is obtained through a reduced total cross-sectional area of holes or reduced thickness of the disc; there are however some limitations related to design, stress and load that the disc must transfer. Thickness h of disc n, for the purpose of the analysis has been identified with y_n, while $\widetilde{y}_n^{(k)}$ is used to identify height to which the material fills the hole in disc n before cut k.

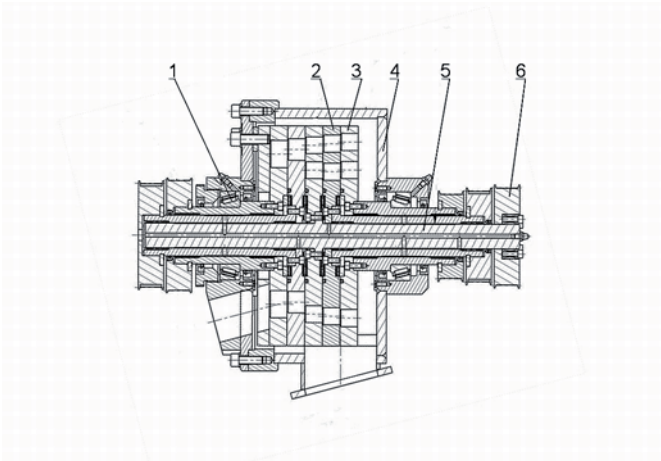


Fig. 3. The working unit of the five-disc, multi-hole RWT-5KZ grinder [4]: 1- bearing, 2- grinding disc, the so-called „preceding" disc, 3- grinding disc, the so-called "next" disc 4- body, 5- shaft, 6- pulley

2. Each point of the cross-sectional area of the holes in the next disc should at some time during the relative motion of the disc be located within the cross-sectional area of the holes in the preceding disc and when the cutting occurs, the so-called next hole of the given layer (of the next disc) is closed for the transfer to the next layer.

The analysis shows that cut k after disc ($n-1$) occurs earlier than cut k after n disc. With such numbering of cuts/quasi-cuts (the numbers begin with the first cut on each disc), the grain may, at each boundary between discs, be subject to a cut identified with the same number. Once a given hole is filled, the common cross-sectional area of the holes begins to decrease and cutting (quasi-cutting) occurs. An

initial assumption was made that each single grain within the cross-sectional area of the grinding holes is subject to cutting. The orientation of grains in relation to the cutting plane is random with even distribution. Grains of any length will be disintegrated with the same probability into two smaller particles, with the total length equal to the length (size) before cutting. The division of the grain as a result of quasi-cutting always occurs in the material located in the preceding disc before filling. With efficient cutting, the distribution of length of grains that filled the empty space in disc (n+1) changes according to the following relation:

$$\tilde{\rho}_{n+1}^{m}(x) = A_{n,m}\rho_n^m = \left(1 - \frac{x}{y_{n+1} - \tilde{y}_{n+1}^{m}}\right)\rho_n^m(x) + \frac{1}{y_{n+1} - \tilde{y}_{n+1}^{m}}\int_x^{l_{\max}} \rho_n^m(l)\,dl, \tag{2}$$

and of those that remained in disc n:

$$\tilde{\rho}_n^{m+1}(x) = \tilde{B}_{n,m}\rho_n^m = \left(1 - \frac{x}{\tilde{y}_n^m}\right)\rho_n^m(x) + \frac{1}{\tilde{y}_n^m}\int_x^{l_{\max}} \rho_n(l)\,dl, \tag{3}$$

where: A, B – stochastic operators for m-th cutting, n-th disc.

For the purpose of simplification, an assumption has been made that after a cut, the distribution of disintegrated grains in hole spaces of disc (n+1) will be even (cut fraction and fraction already in the hole before the cut will mix) and therefore it will constitute the weighted average of ρ_{n+1}^k and ρ_n^k:

$$\rho_{n+1}^m(x) = \frac{\tilde{y}_{n+1}^m}{y_{n+1}}\rho_{n+1}^{m-1} + \frac{y_n - \tilde{y}_{n+1}^m}{y_{n+1}} A_{n,m}\rho_n^m(x). \tag{4}$$

4. Dynamics of efficient movement of grain material

Technological transport of the feed material within the working space of the analysed grinders is not even due to the nature and dynamics of the design and properties of the disintegrated biomass. This analysis begins when planes of two adjacent holes in adjacent working discs overlap and their common cross-sectional area caused by the difference in angular velocity begins to increase. The linear movement of the feed material is caused by the longitudinal force from the feeding screw and effective component of the gravitational force (inclined rotation axis of the disc package in the analysed grinder) (fig. 4). These forces cause ground biomass move between subsequent interconnecting holes in respective structural layers of the machine. Because the main axis of the working unit of the grinder (device) is not vertical, the direction of gravitational force is not perpendicular to the plane of the grinder rotational discs. Angular velocity of a disc results in an additional component of centrifugal acceleration. The sum of those two accelerations within the hole constitutes effective gravitational acceleration. Its direction depends on the phase of disc hole rotation.

When the common area of an overlapping pairs of holes reaches temporary maximum value and begins to approach zero the quasi-cutting process begins. For the purpose of further analysis, the following simplifications were made with respect to the design of working holes: cross-section of a hole is a convex, the volume of the convex is a cylinder (Cartesian product of cross-sectional area of the hole and thickness of the hole), the hole can connect only with one hole in the neighbouring discs at a time. The following structural and technical parameters were determined: thickness of n-th disc, y_n, minimum height (at the disc end) up to which biomaterial fills the hole in n-th disc following k-th cut, $\tilde{y}_n^{(k)}$ volume of grain biomaterial in the hole of n-th disc following m-th cut, V_n^m.

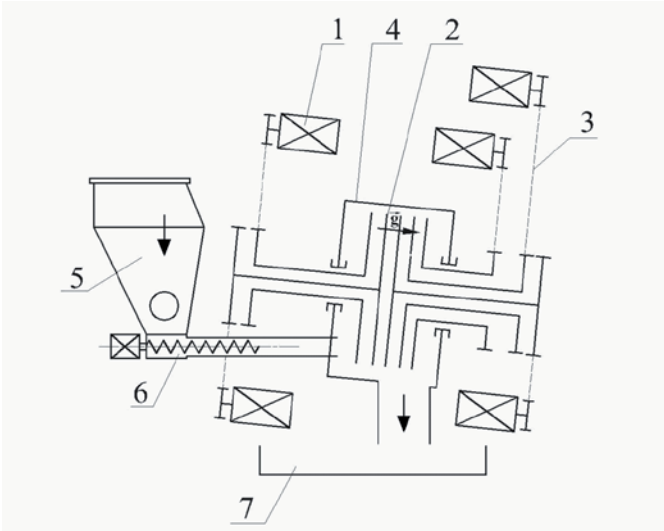


Fig. 4. Kinematical diagram of the multi-disc with inclined axis of working disc rotation for gravitational facilitation of feed movement: 1-electric motor, 2-grinding discs, 3-gears, 4-working space body, 5-container, 6-feeding screw, 7-finished product tank

At this stage, a zero angle of repose for ground medium was assumed (material behaves like a liquid) [7]. For this simplified assumption, the material surface in the working hole, at any time of its movement, is perpendicular to the direction of the effective gravitation vector within the hole area. In addition, an assumption was made that the variability of the effective gravitation within the hole area is omitted and the material surface is a flat surface on the side hole plane. When the common area of the upper and lower hole area is maximal, the common area starts to decrease. Generally, no cutting operation occurs already at this point as there is no need yet for the material to fill the entire common area. Therefore, such phase of the relative motion must be established where the entire common area of the holes cross-sections is filled with the material. The following simplifications are assumed:

1) Each hole is unambiguously assigned its cross-sectional point. This may be e.g. its geometrical centre. The relative motion phase for two holes can be described by providing an angular distance α between the centres of the holes. The ends of the ranges α are marked as α_p, α_k. The D, G, CW functions were defined which return, for a specified $\Delta\alpha$, the area of the following (lower) hole, preceding (upper) hole, and their common area. The symbol α_{max} is used to identify the value of the angular distance for which $|CW|$ function assumes a maximum value (if such maximum is achieved for more than one value α, the highest one of them is assumed as α_{max}). The average vector of the effective gravitation is marked with $\vec{g}_D, \vec{g}_G$ (in the lower and upper hole respectively). A simplified assumption was made that the centrifugal components $\vec{g}_D, \vec{g}_G$ are parallel to each other (they only differ in terms of value as a result of the difference in angular velocities of the two neighbouring discs). V^G, V^D, S^G, S^D stand respectively for the volumes of the preceding and next hole and the cross-sectional areas of the preceding and next hole.

2) For the vector $\vec{g} \in \mathbb{R}^3$ the plane $H_{\vec{g}}$ perpendicular to it was determined, tangent to the edge of the set $CW(\alpha)\times\{0\}$. There are two such planes – one positioned below the set $CW(\alpha)\times\{0\}$ and the second above it (existence of exactly two tangent planes results from the cross-sections common area convexity). The

upper plane was selected. It represents the surface of the material at the point when cutting is started.

3) The cylinder $G(\alpha) \times [0, y_n] \subset \mathbb{R}^3$ was divided with the plane $H_{\vec{g}_G}$. The volume obtained over the plane was identified as (fig. 5):

$$V_{\vec{g}_G}^G(\alpha) = \int_G min\left\{y_n, H_{\vec{g}_G}(r)\right\} d^2r \tag{5}$$

This function attributes to the relative motion phase the volume which remains in the preceding hole at the time when the cutting starts.

4) Similarly, the cylinder $D(\Delta\alpha)\times[-y_{n+1},0]\subset\mathbb{R}^3$ was divided by the plane $H_{\vec{g}}$. The volume obtained under the plane was marked as (fig. 5):

$$V_{\vec{g}D}^D(\alpha) \int^D min\left\{y_{n+1}, H_{\vec{g}_D}(r) + y_{n+1}\right\} d^2 \tag{6}$$

This function attributes to the relative motion phase the volume which remains in the next hole at the time when the cutting starts.

5) The expression $(V_{\vec{g}G}^G + V_{\vec{g}_D}^D)(\alpha)$ for the range $\alpha \in (\alpha_{\max}, \alpha_k)$ is a decreasing function. This function being reversed on its image, the phase of the motion where the cutting starts is obtained for selected volume of material in both grinding holes.

$$\alpha_c(V) = \left(V_{\vec{g}G}^G + V_{\vec{g}D}^D\right)^{-1}(V) \tag{7}$$

This function was extended to include the set $[0, V^G + V^D]$, putting α_k on the left and α_{max} on the right hand side compared with the original function domain.

6) The common area when the cutting for selected volume starts (fig. 5):

$$S_c = \left|\left(Cw \cdot \left(V_{\vec{g}}^G + V_{\vec{g}}^D\right)^{-1}\right)\right| : \left|O, V^G + V^D\right| \to \mathbb{R} \tag{8}$$

This is an increasing function – the highest the amount of the feed material in both working holes, the quicker the cutting is started, involving thus a larger common area of cross-sections.

The resulting function can be treated as the function with the four variables – material volume in both holes, direction of the effective gravitation radial component and the value of the effective gravitation radial components in both holes S_c: $[0, V^G + V^D] \times S^1 \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$. These functions must be determined numerically and treated as the encoding of the influence of the holes shape on the quasi-cutting process for various directions of effective gravitation vectors.

It was assumed that an angle of repose is determined for given material and it does not depend on the distribution of the grain length and that the material is moved only in the direction of the forced displacement (movement direction of the preceding hole in relation to the next hole). This means that the material is inclined down the plane by the angle of repose in a direction corresponding to the direction of the material forced displacement. Because the direction of the relative motion of the holes is always perpendicular to the direction of the centrifugal acceleration, it means that the effective gravitation vector is modified by turning it around the centrifugal component by the angle of repose.

In the cylindrical coordinate system with basis vectors $\{e_r, e_\varphi, e_z\}$, the effective gravitation vectors has the coordinates $(\omega^2 r, 0, g)$. After it is turned around the radial component by the angle γ, the effective gravitation vector will have the components $(\omega^2 r, \pm g\sin\gamma, g\cos\gamma)$. The mark "–" corresponds to the situation where the preceding disc

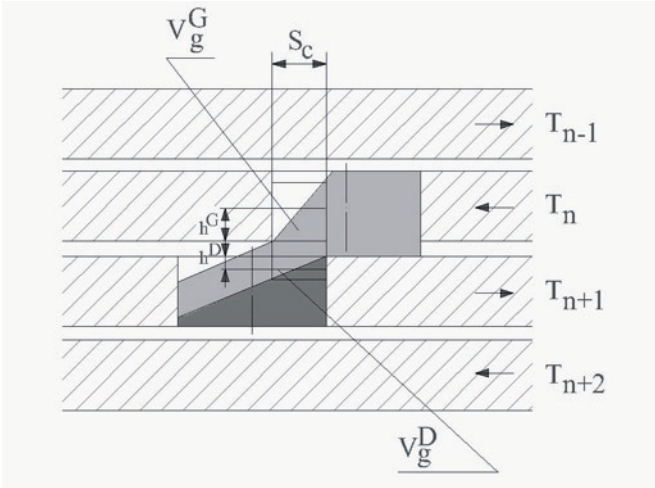


Fig. 5. Diagram presenting two adjacent working holes of the quasi-cutting unit filled with biomass grains. Tn-1 do Tn+2 - next grinding discs, hG – material column height before the cutting plane, hD – material column height behind the cutting plane, VgG- calculated volume of material before the cutting plane, VgD- calculation volume of material behind the cutting plane, SC – cross-sectional area of common pair of quasi-cutting holes

rotates in relation to the next disc with a positive angular velocity and the mark "+" represent the reverse situation.

Cutting is always done in the material which was present in the preceding disc and within the area $S = S_C\left(V_n^m + V_{n+1}^m\right)$ before feeding. The height of the material column over the cutting plane is (fig. 5):

$$h^G = \int_{cw(\alpha_c)} min\left\{y_n, H_{\vec{g}_G}(r)\right\} d^2r \tag{9}$$

and under the cutting plane:

$$h^D = \int_{cw(\alpha_c)} min\left\{y_{n+1}, -H_{\vec{g}_D}(r)\right\} d^2r \tag{10}$$

The probability that a particle from within a length range of $(x, x + dx)$ will be present in the material below the cutting plane following cutting is:

$$\rho_n^m dx - \rho_n^m \frac{x}{h^D} + \int_{l_{\min}}^{l_{\max}} \frac{l}{h_D}\frac{dx}{l}\rho(l)dl \tag{11}$$

The first component corresponds to the probability before the cutting. The probability that a particle with x size reaches the cutting area must be deducted from this component. It is expressed by the occurrence probability product with regard to the occurrence of a particle with x size and the relation x/h^D (particle end must be within the segment $(0, x) \subset (0, h^D)$). The third factor are the probabilities, integrated from x to l_{max}, that there is a particle with a length of $l(\rho(l)dl)$ and that it will reach the cutting plane l/h^D, and that a particle from within a range of $(x, x + dx)$ (dx/l) will be obtained after cutting.

Based on the above, the distribution of the grain length in the material filling an empty space within the disc $(n + 1)$-th changes in the following way:

$$\tilde{\rho}_{n+1}^m(x) = A_{n,m}\rho_n^m = \left(1 - \frac{x}{h^D}\right)\rho_n^m(x) + \frac{1}{h^D}\int_x^{l_{\max}} \rho_n^m(l)dl \tag{12}$$

while in the material remaining in the n-th disc (analogous reasoning):

$$\tilde{\rho}_n^{m+1}(x) = \tilde{B}_{n,m}\rho_n^m = \left(1 - \frac{x}{h^D}\right)\rho_n^m(x) + \frac{1}{h^G}\int_x^{l_{\max}} \rho_n(l)dl \tag{13}$$

It must be remembered that the column of the material subject to cutting is not on the whole the entirety of the material which has been moved to the lower hole. Its volume is $S_c \cdot h^D$, while that of the entire material moved from the preceding hole to the next one is:

$$V_{\bar{g}}^{D}(\alpha_c) - V_n^m$$

This means that the second and third component in (13) should be multiplied by their volumetric ratio:

$$\tilde{\rho}_n^{m+1}(x) = \tilde{B}_{n,m}\rho_n^m = \left(1 - \frac{Sc \cdot x}{V_{\bar{g}}^{D}(\alpha_c) - V_{n+1}^m}\right)\rho_n^m(x) + \frac{Sc}{V_{\bar{g}}^{D}(\alpha_c) - V_{n+1}^m}\int_x^{l_{\max}} \rho_n(l)\,dl \tag{14}$$

It was assumed for simplification purposes that after cutting the distribution of the grain length in the disc (n+1)-th will be even (cut fraction and fraction already in the hole before the cut will mix) and therefore it will constitute the weighted average of $\rho_{n+1}^{(m-1)}$ and $\tilde{\rho}_{n+1}^{m-1}$:

$$\rho_{n+1}^m(x) = \frac{V_{n+1}^m}{V^D}\rho_{n+1}^{m-1} + \frac{V^D - V_{n+1}^m}{V^D}B_{n,m}\rho_n^m(x) \tag{15}$$

The filling of the quasi-cutting unit and thus the cutting efficiency depends on the value of the function V^D, V^G and S_c, which depend on the direction of effective gravitation and the sum of material volume in both holes before cutting $(V_{n+1}^m + V_n^m)$.

After grains are cut, two layers of the feed material move in relation to each other according to the rotational direction of the neighbouring discs and the gradient of mutual velocities. Material particles are removed from the preceding hole (they are subjected to the effect of the gravitational force effective component with a direction perpendicular to the inter-disc gap), but they are not removed from the next hole.

After cutting, length distribution will be:

$$\tilde{\rho}_n^{m+1}(x) = B_{n,m}\rho_{n,m} = \begin{cases} \tilde{\rho}_n^{m+1}(x)\left(\int_{l_{\min}}^{l_{\max}} \tilde{\rho}_n^{m+1}(x)\right)^{-1} & x \rangle l_{\max} \\ 0 & x \langle l_{\min} \end{cases} \tag{16}$$

The level of material for m-th cut (before m+1 cut) in n-th gap, $\tilde{y}_n^{m+1}$ is:

$$\tilde{y}_n^{m+1} = \left(y_n - y_{n+1} + \tilde{y}_{n+1}^k\right)\left(1 - \frac{\int_0^{l_{\min}} \tilde{B}_{n,m}\rho_n^m(x)\,x\,dx}{\int_0^{l_{\max}} \tilde{B}_{n,m}\rho_n^m(x)\,x\,dx}\right) \tag{17}$$

To obtain distribution across the entire space of the hole before (m+1)-th cut (after re-feeding), the weighted average must be determined:

$$\rho_n^{m+1}(x) = \frac{y_n - \tilde{y}_n^{m+1}}{y_n}A_{n-1,m}\rho_{n-1}^m + \frac{\tilde{y}_n^{m+1}}{y_n}B_{n,m}\rho_n^m \tag{18}$$

The operator $B_{n,m}$ is no longer a linear operator like $A_{n,m}$ because it depends on the level of the material left in the *n-th* after *m-th* cut: $\tilde{y}_n^{m+1}$ and is the function of the probability distribution in the material (which determines what part of the material is removed from the machine during the cutting). So that $B_{n,m}$ can be treated as linear operators, the values y_n^m should be treated at every stage of the procedure as predetermined and iterative, which take into consideration the results of tests and experiments.

Determination of the cross-sectional area of the common part as the cross-section of grinding must be based on efficient mathematical procedures using the essence of the integral calculus or analytical geometry. Based on comprehensive analyses of cross-sections and multi-disc grinding resistances, it can be stated that the method for calculation of the grinding cross-section common area, in the analyses of the multi-disc seed grinding energy efficiency, depends directly on computing capabilities [8].

During the modelling of the grinding surface and the common surface, integration of the momentary grinding cross-section was used at the first stage (fig.1 and fig.6):

$$S_C = \int_{x_1}^{x_2}\left\{b_2 + \left[R_2^2 - (x - a_2)^2\right]^{1/2}\right\}dx - \int_{x_1}^{x_2}\left\{b_1 - \left[R_1^2 - (x - a_1)^2\right]^{1/2}\right\}dx \tag{19}$$

where:

a_1, a_2, b_1, b_2 - holes coordinates: C_1 i C_2,
R_1, R_2 – holes radius

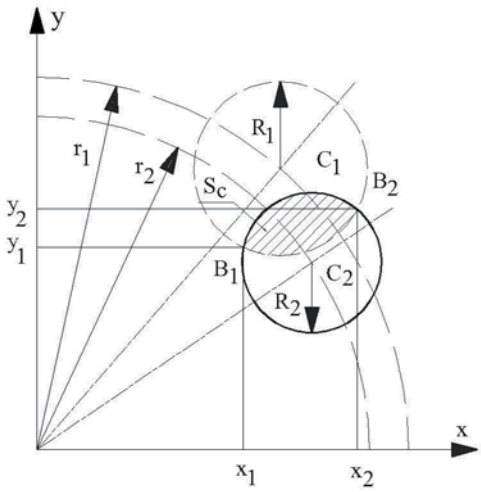


Fig. 6. Common area of two holes, grinding cross-section between the edges of neighbouring grinding holes: C1, C2 - working holes centres, R_1, R_2 – working holes radiuses, r1, r2 – working holes positioning radiuses, x1, x2, y1, y2 - coordinates for B1 and B2 quasi-cutting pair circles intersection points, SC –quasi-cutting hole pair common area

5. Summary

To sum up, it must be remembered while identifying the mathematical model of the grinder structure solution that the mathematical model describes the actual grinder only in an approximate way. There may be a number of various discrepancies between the model and the grinder which most frequently occur for the following reasons:

1. For derivation of the mathematical dependencies, classic mechanics laws are usually used which were formulated in reference to stiff non-deformable bodies. Moreover, various heat, friction and wear related phenomena as well as other phenomena are omitted as a rule which accompany the analysed mechanical processes.
2. The discrete model being identified has a finite number of degrees of freedom, while the mechanical grinder is a system with an infinite number of degrees of freedom.
3. The mechanical multi-drive grinder, within the full range of blades dynamic work and two overlapping holes, is mostly a nonlinear system. However, within some ranges of the tool common area (pairs of holes), the grinder behaves like a linear mechanical system, and therefore the assumption of the linear model is fully justified in reference to these ranges.

Furthermore, it must be added that the conditions in which the identification of the solution model is carried out may be materially different from the actual operation conditions of the grinder. This refers mainly to the nature of signals (determined or random) and the points of forcing signals. The mathematical model is in such cases defined under conditions differing from the actual operating condi-

tions of the grinder, which may result in significant features of the mechanical system being omitted. If, for example, the ranges of the actual forcing signals are much wider than the ranges of forcing signals conducted at the time of a planned experiment, certain nonlinearities can be found important for the operation of the innovative system.

Omission of certain qualitative features of forcing signals may result in the distortion of the model actual dynamic properties. Bearing that in mind, one should, during innovative implementation, inspect carefully dynamic properties of the grinder and such analysis should demonstrate that the properties of the grinder correspond to its mathematical model.

6. Conclusions

Phenomena, processes and relations in multi-disc grinding, despite their complexity, are among those that can be described formally. The answers to the questions about process factors (activities and methods), structural features (methods, equipment and systems), operating conditions and their influence on the dynamics and efficiency of biomass grain grinding, based on the example of a multi-disc grinder, was possible if assuming infinitely quick propagation of quasi-cutting stresses within the scope of:

- probability distribution for a grain length,
- probability distribution for particles in the stream leaving the machine through the gap between the discs.

At m-th cut, the output stream of particles leaving the machine through the inter-disc gap between n-th and (n+1)-th disc is described by a probability distribution:

$$s_n^m(x) = \begin{cases} \int_x^{l_{\max}} \rho_n(l)\,dl \left(\int_0^{l_{\min}} \int_x^{l_{\max}} \rho_n(l)\,dl\,dx \right)^{-1} & l \langle l_{\min} \\ 0, & l \rangle l_{\min} \end{cases} \tag{20}$$

and its volume equals:

$$V = \left(y_n - y_{n+1} + \tilde{y}_{n+1}^k \right) \frac{\int_0^{l_{\min}} \tilde{B}_{n,m} \rho_n^m(x)\,x\,dx}{\int_0^{l_{\max}} \tilde{B}_{n,m} \rho_n^m(x)\,x\,dx} \tag{21}$$

The effective component of gravitational force and the angle of repose influence the dynamics and efficiency of the five-disc multi-hole grain biomass RWT-5KZ grinder. The significance of the influence is visible in their relation to basic parameters of the multi-disc grinding process, e.g. surface filling (crosswise) of the carrying and grinding space (operating space), actual volume and weight (mass efficiency) of the multi-edge grain grinder hole space filling. It was described with the variability of the operating space structural features, brought to the volume V^G, V^D, and the cross-sectional area S^G, S^D of the co-working pairs of the quasi-cutting holes, kinematical parameters of motion and the influence of the vector of effective components of gravitational force ($\vec{g}_G$, $\vec{g}_G$) caused by the inclination of the main axis of the operating discs package.

The research financed by the funds of the National Centre for Research and Development in 2010/2013 as a development project.

References

1. Chwarścianek F. The construction all increase of crumbling effectiveness. The Archive of Mechanical Engineering, 2007, Vol. LIV, No 4: 391÷408.
2. Detyna J. Analysis of nonequilibrium stases in the sieve separation process. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2011, 1(49): 78-85.
3. Flizikowski J. The Construction of the Food Grinders. Wydawnictwo Uczelniane Akademii Techniczno-Rolniczej w Bydgoszczy 2005.
4. Flizikowski J. Doskonalenie badań i rozwoju rozdrabniaczy. Inżynieria i Aparatura Chemiczna, 2006, 1-2: 38-39.
5. Flizikowski J, Bieliński M. Multidisc Grinder Especially for Grains. Patent.RP-144 566.
6. Flizikowski J, Lis A. Optymalizacja rozdrabniacza wielotarczowego. Inżynieria i Aparatura Chemiczna 2007, 46(38), Nr 1: 50-52.
7. Khazaei J, Ghanbari S. New method for simultaneously measuring the angles of repose and frictional properties of wheat grains. International Agrophysics, 2010, 24: 275-286.
8. Macko M, Czerniak J. The evolutionary method for optimising disk design of multi-edge grinders. Journal of Theoretical and Applied Mechanics 8211, 08/2011.
9. Razavi S. M. A, Farahmandfar R. Effect of hulling and milling on the physical properties of rice grains. International Agrophysics, 2008, 22: 353-359.
10. Tomporowski A. Studium efektywności napędu i rozwiązań innowacyjnych konstrukcji wielotarczowych rozdrabniaczy ziaren biomasy. Societsa Scientiarum Lublineneis, 2011.
11. Ulrich S, Schroter M, Swinney H. Influence of friction on granular segregation. Physical Review, E, 76, 042301, 2007.
12. Walton O. Effects of interparticle friction and particle shape on dynamic angles of repose via particle-dynamics simulation. Proc. Conf. Mechanics and Statistical Physics of Particulate Materials, June 8-10, La Jolla, CA, USA, 1994.

Dr inż. Andrzej TOMPOROWSKI
Department of Mechanical Engineering
University of Technology and Life Sciences in Bydgoszcz
ul. Prof. S. Kaliskiego 7, 85-789 Bydgoszcz, Poland
E-mail: a.tomporowski@utp.edu.pl

Prof. dr hab. inż. Marek OPIELAK
Lublin University of Technology
ul. Nadbystrzycka 38, 20-618 Lublin, Poland
E-mail: m.opielak@pollub.pl

Article citation info:
ORKISZ M, ŚWIĘCH Ł, ZACHARZEWSKI J. Fatigue tests of motor glider wing's composite spar. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 228-232.

Marek ORKISZ
Łukasz ŚWIĘCH
Jan ZACHARZEWSKI

FATIGUE TESTS OF MOTOR GLIDER WING'S COMPOSITE SPAR

BADANIA ZMĘCZENIOWE KOMPOZYTOWEGO DŹWIGARA SKRZYDŁA MOTOSZYBOWCA

The paper presents experimental and numerical investigation of wing's spar. Part of the spar was subjected to one step fatigue test, covering 10,000 load cycles corresponding to the oscillations of the load factor from nz min = −3.7 to nz max = 5.7. Such test is proposed as an alternative to the full loading spectrum tests. During the experiment 3D scan was used to rapid inspection of sensitive structure's areas. Application of optical strain gauges based on a fiber Bragg's grating allowed to observe the phenomenon of local, periodical strengthening of the structure.

Keywords: *composite structures, fatigue tests, FBG strain gages, finite element method (FEM), 3D scanning.*

W pracy przedstawiono badania eksperymentalne i numeryczne dźwigara skrzydła. Fragment dźwigara poddano jednostopniowemu testowi zmęczeniowemu, obejmującemu 10.000 cykli obciążeń odpowiadających oscylacjom współczynnika obciążeń od nz min= −3.7 do nz max= 5.7. Test taki proponowany jest jako alternatywa dla próby z zastosowaniem pełnego spektrum obciążeń. W trakcie badań wykorzystano skanowanie przestrzenne, jako propozycję szybkiej metody inspekcji newralgicznych obszarów konstrukcji. Zastosowanie do pomiaru odkształceń systemu światłowodowych czujników tensometrycznych opartych na siatce Bragga'a pozwoliło na zaobserwowanie zjawiska lokalnego, okresowego umacniania struktury.

Słowa kluczowe:*struktury kompozytowe, badania zmęczeniowe, tensometry oparte o siatkę Bragg'a, metoda elementów skończonych (MES), skanowanie przestrzenne.*

1. Introduction

The production started of new motor glider type requires fatigue tests to prove real fatigue life of such structure. Modern constructions of composite gliders and motor gliders should fulfill minimum of 9,000 flight hours. It gives the possibility of more than two decades of service at the annual statistics of the average user flying time. Such conditions have been the subject of our research work [1,2,3,5,6,7,8]. Safe life of composite thin-walled structure means that operating loads in time of service do not cause such weakness of the structure, which violates the applicable safety factor (the ratio of loads to the destructive limit loads). For all aircraft structures, this factor has a value of 1.5. The remained strength at the end using of such structure cannot be less than 150% of allowable loads.

Evidence of fatigue life can be carried out by the fatigue tests. Fatigue testing programs are associated with the various features of the fatigue properties of composite structures. Fiber-reinforced composites feature is that it is insensitive to variable loads with low values, but their fatigue life and remained strength is significantly reduced as a result of single loads closed to the lower limit of scatter of immediate strength.

The above-mentioned properties of composites mean that instead of expensive whole motor glider bench studies could be more economical to show the fatigue life of some parts of the airframe structure. Rational will be to divide the fatigue test of a composite structure to the following steps:

a. Studies of force introduction nodes;
b. Studies of most loaded structure element – wing's spar;
c. Studies of wing-fuselage unit.

2. Loadnig and attachment conditions of tested spar

The main task during the preparation of experimental studies of actual structures is to ensure conditions of loading and attachment as accurately as possible similar to the working conditions of the element. The tested composite spar is installed in a motor glider fuselage by two fittings in its bayonet part.

During the motor glider's flight, as a result of pressure distribution on the surface of the wing appear forces, which are transmitted to the spar through the structure of the wing. Thus, spars works in a complex load state. Such conditions are extremely difficult to reproduce during fatigue testing, due to the need to control many parameters, which leads to a complex loading system.

For simplicity, it was decided to carry out a one-parameter fatigue test. For the wing spar is assumed that shear force is the most influencing load. This way, occurrence of the shear force and bending moment on the bayonet part was provided with values corresponding to real load of the structure.

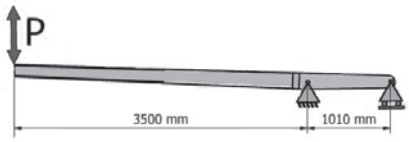


Fig. 1. Scheme of support and loading of spar

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

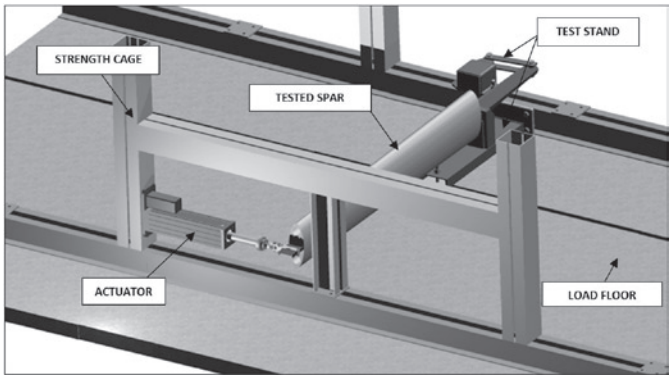


Fig. 2. CAD model of test stand

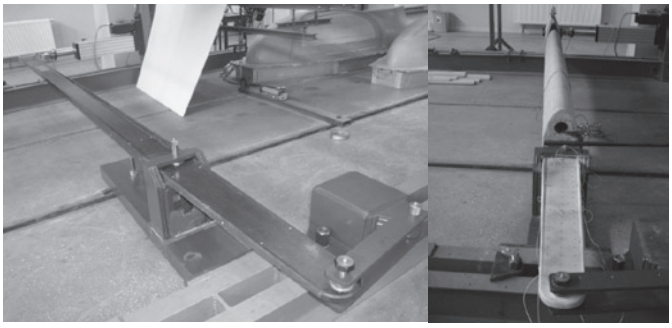


Fig. 3. General view oftest standwith attachedspar

Due to the dimensions of the spar (4.5 meters), CAD model of complete test stand with a part of the strength cage and load floor which are the place of attachment, was designed (fig.2).

In order to protect against the loss of stability of the spar, torsion box was made of glass-fiber reinforced composite intensified by wooden ribs. Torsion box task is to take over the torque, which was created by twisting of the spar.

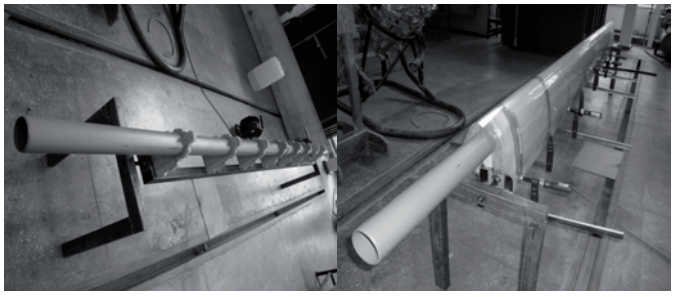


Fig. 4. Torsion box

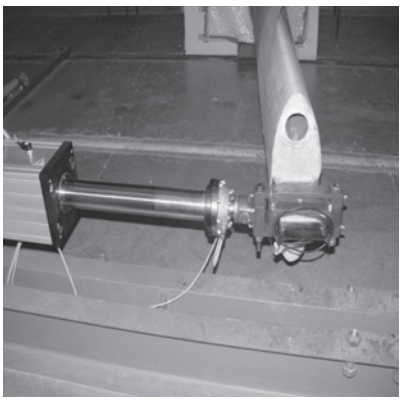


Fig. 5. Force introduction onspar

3. Load spectrum

Using the experience and the suggestions made in [1], the test spar was subjected to one step load spectrum, what avoids the need for long-lasting and expensive research on the full, operational spectrum.

During the test, the power was generated by electro-mechanical Zwick-Roell cylinder, with force level control. Figure 6 shows carried load spectrum. Cycle asymmetry factor was R= −0.65 which corresponds to changes in load factor in the range of n = 5.7 to n = −3.7. Test program established to carry out 10,000 loadcycles.

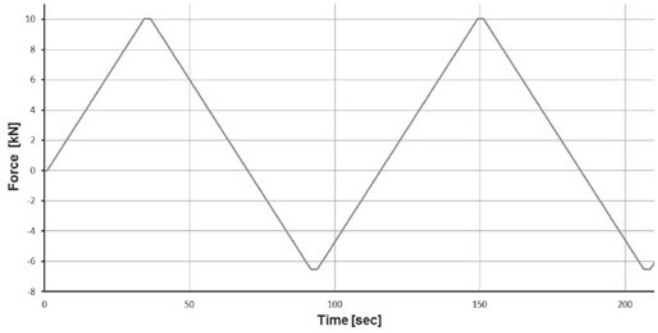


Fig. 6. Realized load spectrum

4. Spar construction

Tested spar was a fully composite structure, constructed as an I-beam. The most sensitive area is called bracket's part of the spar and connects the wing to the fuselage. The remaining part is responsible for ensuring transfer of shear forces and bending moment from the wing to the fuselage.

Bayonet part has a sandwich structure. Core is made of polyurethane foam DIV-60 sandwiched between twenty-six layers of carbon fabric SGL KDK 8042. In the most loaded areas containing fittings, different layers were used. The whole structure was covered with one layer of glass fabric Interglas 92110. Spar flanges are made of unidirectional carbon composite (roving) Torayca T700G. Wall of the spar inside the wing creates a laminar structure with properly graded carbon fabrics.

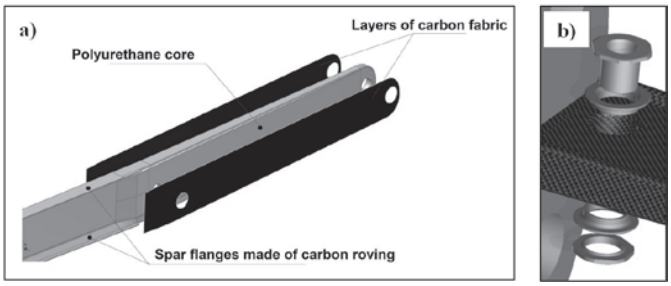


Fig. 7. a) Scheme of composite structures, b) Construction of the labyrinth-lock

Support points (force introduction on the composite structure) were made as a steel sleeve, forming the so-called labyrinth lock. A scheme of such structure was shown in figure 7b.

The spar was designed and manufactured in the Aeronautical Department of Warsaw University of Technology.

5. Finite element analysis of spar

The aim of numerical analysiswasto complete thefatigue tests. This allowedto visualize fieldsof displacements andstressesinthe testedstructure. Calculationswere performedwith MSCNastran2010, usingproceduresof linearstatic analysis.

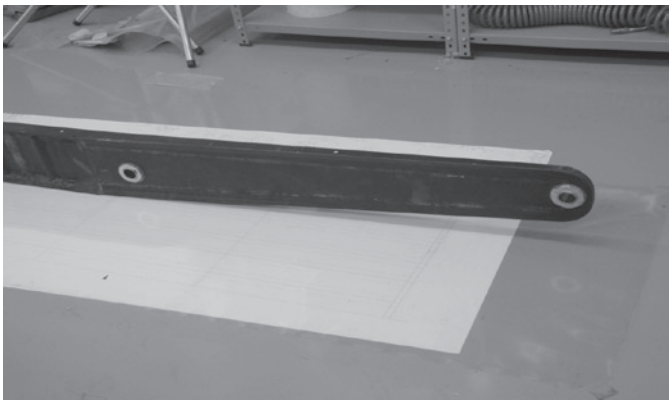


Fig. 8. Bayonet partof the testedspar

Fig. 9. Simplified geometry of thespar with support markers

On the basis ofthe detailedgeometry, simplifiedsolid model of spar was made (fig. 9), which was used to perform calculations.

Due to the complex spar construction (sweep and dihedral), it was decided to made the whole tested piece of the spar, what eliminated the necessity of applying loads as the internal forces. Method of attachment and load corresponds to the experimental conditions.

FEM model was built of both elements of 1D, 2D and 3D. Spar flanges and foam of the bayonet part were made of eight-node solid elements, with corresponding material properties, limited to isotropic, elastic model.Carbon fabric surrounding the core and the remainder of the wall sheathing simulate four-node plane elements, with orthotropic material model with definition of all laminate layers. In order to avoid troublesome phenomenon of contact, the steel fittings replaced with rigid 1D elements (fig. 10).

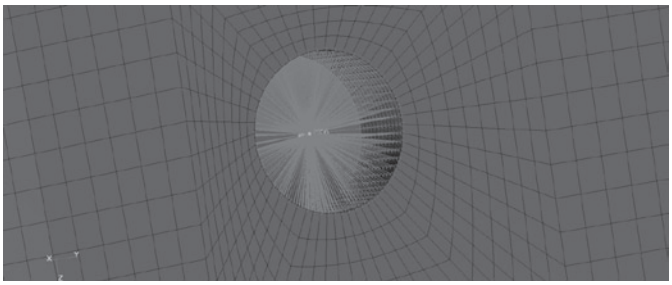


Fig. 10. Part of FEAmesh and 1D elementssimulatingfittings

Due to the low torsional rigidity of spar and to ensure one plane bending, it was taken the possibility of lateral movements of the elements forming the upper and lower spar flanges outside the bracket's part.

Load as during the experiment was carried out by applying shear force to the spar wall. There were performed two series of numerical analysis, corresponding to load factor

n = 5.7 and n = −3.7.

6. Measurement of spar deflection on bracket's part

Fatigue damage can be manifested through unstiffening of composite structure during the operation (tests). Easily measurable parameter of proving the change of stiffness of the structure, is to measure the deflection line of spar bracket's part. During the fatigue test, measurements of deflection between spar supports were made with use of inductive sensor. Data recording was made by using of HBM's acquisition system Spider8.

Fig. 11. Attachment ofthe inductive sensor

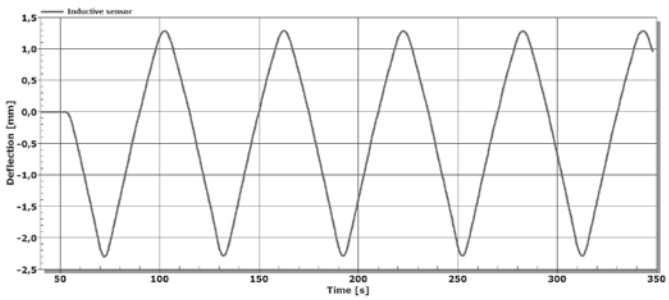


Fig. 12. Samplespardeflections

The size of deflections recorded by the inductive sensor are in the range from −2.3 mm for load n = 5.7 to 1.3 mm for load n = −3.7. Almost identical results were obtained in the FEM analysis (figure 13 and figure 14).

During program realization there was not noticed the change of the value of deflection, which shows no change in stiffness of the tested structure (fig. 15).

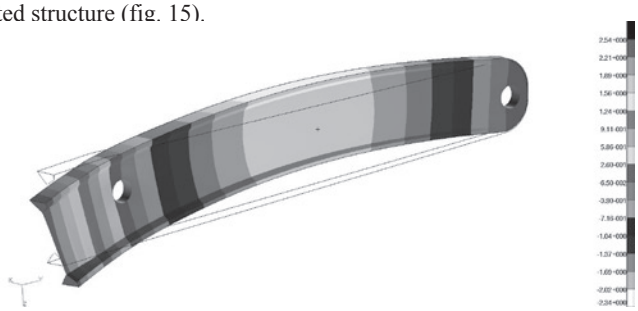


Fig. 13. Spar deflection. Result ofFEM analysis. Loadn=5.7

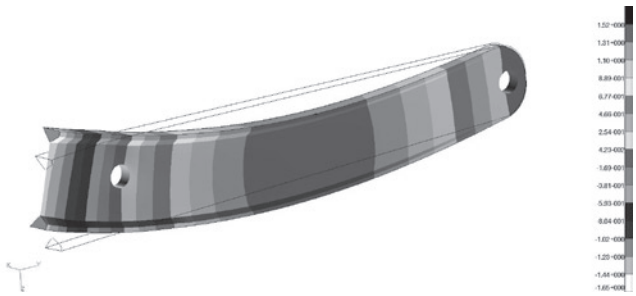


Fig. 14. Spar deflection. Result ofFEM analysis. Loadn=−3.7

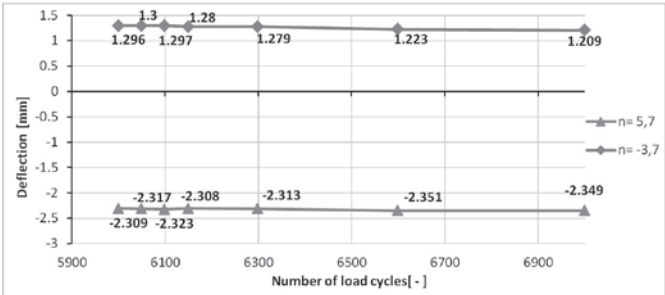


Fig. 15. Changes of deflection during the test

7. Measurement of fitting displacements

Force introduction nodes, caused high levels of stress in acomposite structure in area of fittings (fig.16), which creates arisk of irreversible deformations. One of the alarming symptoms of the possibility of these phenomena is the change of distance between mentioned nodes.

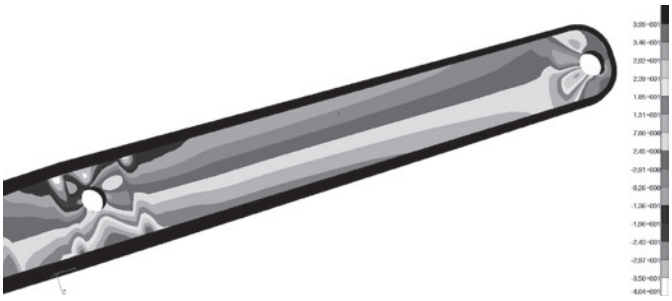


Fig. 16. Stressesin thebayonet part of spar

For measuring the displacement of nodes was used the optical white light scanner ATOS. It allowed three-dimensional digitization of the object, using the shadow moiré method and stereoscopic photography. With applied measuring field the accuracy of 0.01 mm was obtained. Inspections were made after the completion of each 1,000 load cycles.

Fig. 17. ATOS scanner

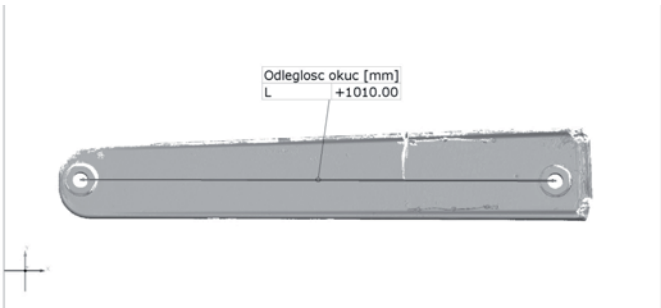


Fig. 18. Spar scanafter2,000load cycles

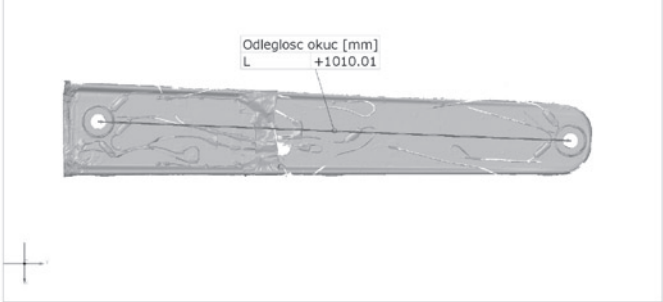


Fig. 19. Spar scanafter5,000load cycles

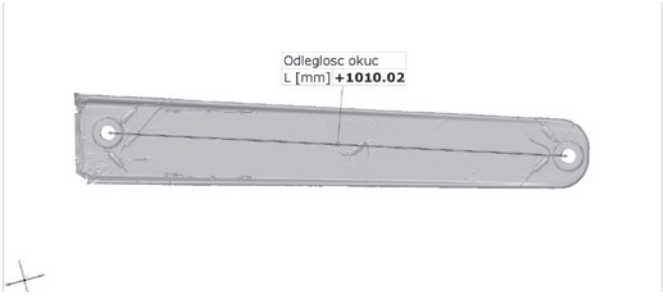


Fig. 20. Spar scanafter10,000load cycles

Inserted images show that the distance between nodes did not change significantly after 10,000 load cycles.

8. Optical strain gages

Experiments with composite spar became an opportunity to test optical strain gage based on fiber Bragg gratings, in conditions of sustained cyclic loading. Compared to the traditional measurement of deformation based on resistance strain gages, tested system presents many advantages, of which the foreground (in aerospace applications) extends its small mass. Reduced weight of such measurement system is linked to the possibility of placing up to 13 points on a single fiber optic fiber [4].

The operation principle of optical strain gage is based on Bragg gratings, in the form of periodic cuts on fiber, with the task of reflecting a specific wavelength of light. The length of the applied grid is thereby

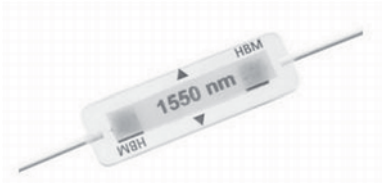


Fig. 21. Illustration of optical strain gage [HBM GmbH]

the base of strain gage (in used sensors, it is about 6 mm).

Signal in the form of a light wave is generated in the device called an interrogator, which is also the receiver of the reflected light.. After reflection, the remainder of the light wave passes through the fiber remains (fig. 22) and is used as a measuring signal for subsequent configuration of sensors with different notched grid.

Measurement of deformations is realized by comparing the reflected light wave length changed by elongation or shortening of the sensor, to the reference wavelength of unstrained sensor.

During the researches, changes in strain near the front ferrule observed (fig. 24). The initial phase (up to 5,000 cycles) showed no significant changes in work of structure. Next load cycles (5,000 – 6,000) resulted the significant decrease in the measured strain, which can indicate the strengthening of the composite structure, associated with the setting of the carbon fibers in a matrix of epoxy resin. Then, over the 1,000 cycles indication returned to baseline. The last phase of research showed the tendency to stiffening structure part with introduction point

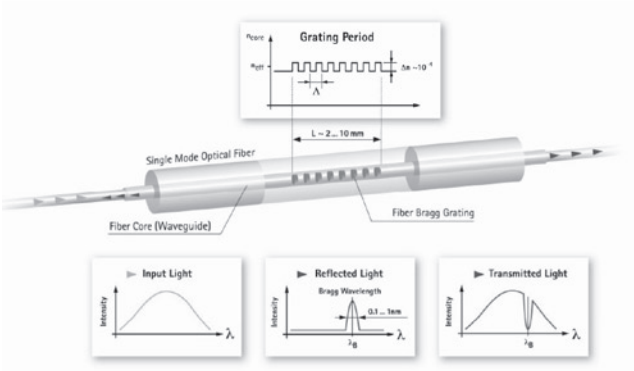


Fig. 22. Principle of Bragg gratings [HBM GmbH]

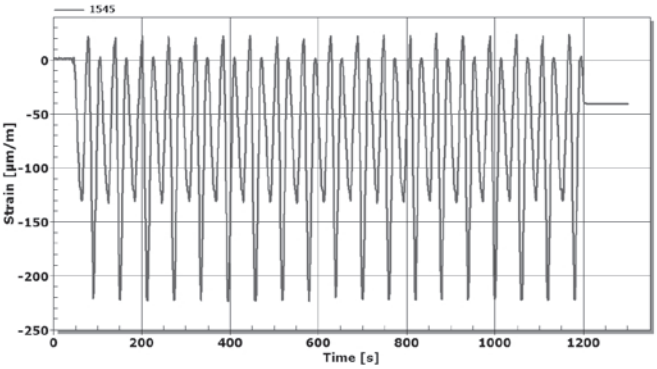


Fig. 23. Example of thestrainhistory ofstrain gagenumber1545(around the frontfitting) -6401-6450cycle

of concentrated forces. It should be noted, however, the fact that the construction in terms of macroscopic demonstrated no significant changes in stiffness, for example, a change that results in the deflection of the beam which demonstrate that the observed phenomenon is local.

9. Inspection of node's geometry

After end of test program, there were no fatigue damages of the spar. Composite parts in the area of primary fitting was subjected to detailed inspection, where also was not observed this kind of changes (fig. 25).

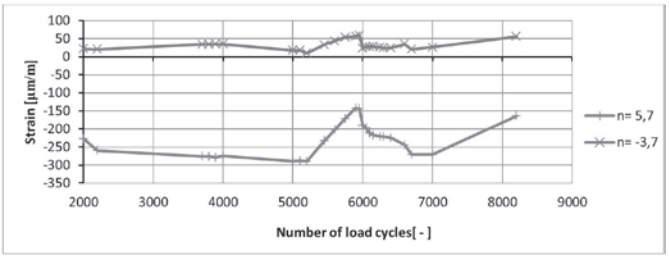


Fig. 24. Changing ofstrainduring the test(strain gage number.1545)

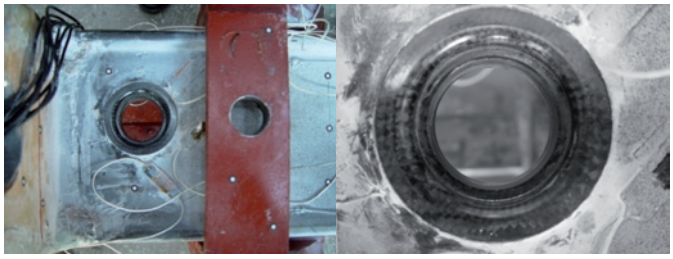


Fig. 25. Structure state after test program

10.Conclusions

It was realized full test program and significant changes as well in the composite parts as in the rigidity of the structure were not observed. It has been proved this way the aimed spar fatigue life.

Reconstruction of studied structures surfaces using 3D scanning avoids the need to carry special equipment, which provides to perform accurate and simple measurements.

First experience with the use of optical strain gages based on fiber Bragg gratings in fatigue tests of composite structures, with high levels of strain, confirm the usability of this type of measurement system. Application of such strain gages allowed to observe the phenomenon of local, periodical strengthening of the structure during the experiment.

Results of fatigue test presented in this paper did not finish the intended research. The presented spar will be subjected to further load cycles, which will lead to the determination of the limited cycle-numbers corresponded to the actual safe life time of tested composite structure.

References

1. Gassner E, Shutz M. The Significanse of Constant Load Amplitude Tests for the Fatique Evaluation of Aircraft Structures. Amsterdam: Proceedings, Symposium on Fatigue Testing of Aircraft Structures, 1959.
2. KenscheCh W. Lifetime of GFRP in a shear web and in the girder of a sailplane wing spar. Technical soaring 2002; 26(2): 51-55.
3. KenscheCh W. Proposal for a certification procedure of extended sailplane lifetime. Technical soaring 2002; 26(2): 32-43.
4. Maul J, Kipp T. Sensing of surface strain with fexible fiber Bragg strain gages. Darmstad: HBM GmbH.
5. Rodzewicz M. Problematyka modelowania w badaniach trwałości zmęczeniowej węzła kadłubowego połączenia skrzydło-kadłub szybowca PW-5. Warszawa: Wydawnictwo PTMTiS, Materiały konferencyjce „Mechanika w lotnictwie", 1994; 473-485.
6. Rodzewicz M. Investigation of the glider load spectra. Technical soaring 2007; 31(1): 2-12.
7. Rodzewicz M. Spektra obciążeń i trwałość zmęczeniowa struktury nośnej szybowców kompozytowych. Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej, 2008.
8. Rodzewicz M, Przekop A. Experimental investigation of the load spectrum and fatigue tests of the PW-5 World Class Glider. Technical soaring 2000; 24: 15.

Professor Marek ORKISZ, Ph.D., D.Sc. (Eng.)
Łukasz ŚWIĘCH M.Sc. (Eng.)
Jan ZACHARZEWSKI, Ph.D. (Eng.)
Deptartment of Aircraft and Aircraft Engines
Rzeszow University of Technology
al. Powstańców Warszawy 8, 35-959 Rzeszów
e-mail: mareko@prz.edu.pl, lukasz.swiech@gmail.com, jzacharz@prz.rzeszow.pl

Article citation info:
JURCA V, ALES Z. Maintenance management systems in agricultural companies in the Czech Republic. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 233-238.

Vladimir JURCA
Zdenek ALES

MAINTENANCE MANAGEMENT SYSTEMS IN AGRICULTURAL COMPANIES IN THE CZECH REPUBLIC

SYSTEMY ZARZĄDZANIA UTRZYMANIEM RUCHU W PRZEDSIĘBIORSTWACH ROLNYCH W REPUBLICE CZESKIEJ

The purpose of this paper is to describe different maintenance management systems being used by Czech companies with a view to the agro-industrial complex in recent years. The paper focuses on maintenance management systems supported by information system (IS) and their implementation in companies. The use of IS helps to create a variety of scenarios with an emphasis on the need for analytical instruments processing data integration in order to improve the efficiency of the maintenance system and identify its weaknesses. While the maintenance management information system (MMIS) is widespread within the agro-industrial scope in the manufacturing industry, in primary agricultural production and agricultural machinery servicing, it has yet to be utilized. The paper also describes some possible applications of MMIS in the sphere of the agro-industrial complex, the differences between their applications and the potential use of MMIS in the case of fully or partially outsourced agricultural equipment maintenance.

***Keywords**: maintenance management systems, information system, preventive maintenance, maintenance costs.*

Przedstawiona praca ma na celu opis różnych systemów zarządzania utrzymaniem ruchu stosowanych w ostatnich latach przez czeskie przedsiębiorstwa ze szczególnym uwzględnieniem zakładów rolno-przemysłowych. W pracy główną uwagę skupiono na systemach zarządzania utrzymaniem ruchu wspieranych przez systemy informatyczne (IS) oraz na problemie ich wdrażania w przedsiębiorstwach. Użycie systemu informatycznego pomaga w tworzeniu różnych scenariuszy, przy czym nacisk kładzie się na zapotrzebowanie na instrumenty analityczne służące do integracji przetwarzanych danych w celu polepszenia wydajności systemu utrzymania ruchu oraz rozpoznawania jego słabych punktów. Podczas gdy użycie systemów informatycznych wspomagających zarządzanie utrzymaniem ruchu (ang. maintenance management information system, MMIS) jest powszechne w zakładach rolno-przemysłowych przemysłu wytwórczego, podstawowa produkcja rolna oraz obsługa maszyn rolniczych czekają jeszcze na ich wdrożenie. W artykule opisano także możliwe zastosowania MMIS w zakładach rolno-spożywczych oraz różnice w aplikacji i potencjalnym wykorzystaniu MMIS w przypadku prac w zakresie utrzymania ruchu maszyn i urządzeń rolniczych całkowicie lub częściowo zleconych firmom zewnętrznym.

***Słowa kluczowe**: system zarządzania utrzymaniem ruchu, system informatyczny, obsługa zapobiegawcza, koszty obsługi.*

1. Introduction

The maintenance status of Czech companies has changed considerably in recent years in favour of maintenance, which previously was often seen by management as a kind of necessary evil, which only draw from corporate resources. Today, in prosperous companies maintenance has become one of integral processes, where it is necessary to apply similar principles of quality management, such as in manufacturing and all other areas of business management. One of the important principles of quality management is a "continual improvement", which has to be a sustainable objective of any organization. [2]

Successful quality management in general, including the maintenance management system is based on full documentation of all related activities. Anyone who deals with maintenance issues is aware that well-functioned efficient maintenance system must be planned and transparently documented. It must be clearly determined when, who and how to carry out the maintenance. It has to be back indictable how, by whom and when the maintenance was performed, at what costs was the maintenance performed and what is the quality and reliability of the whole system and its elements. There are many indicators which should be documented and frequently analyzed and therefore for the most companies it is necessary to use the information system. Maintenance management information system (MMIS) is designed to simplify and standardize documentation processes of maintenance data. Beside that such an information system promotes the process approach and system approach to management which enables a continuous improvement of maintenance. [5] Maintenance information system in particular allows decision making based on facts, which is accessible in the system and which promptly provides information in order to make a decision on the rational basis.

The MMIS use is already widespread in the Czech manufacturing industry. The MMIS was first introduced in the mid 90's in automotive and related sectors. At the turn of the century the MMIS began to be utilized also in agro-industrial complex (dairies, bakeries, sugarhouses, breweries etc.), but in the enterprises with primary agricultural production the MMIS has been used only exceptionally although the agricultural machinery is getting to be more computerized and more expensive in terms of their purchase and operation.

A well-functioning maintenance system can significantly reduce operating costs, particularly in the following areas:

- Enhance the reliability of production equipment and thus improve its elementary properties (especially availability, safety and durability) [7]
- Significantly change the ratio between maintenances after failure and planned maintenances in favour of planned mainte-

nance and thus reduce downtimes and subsequently increase machine utilization
- Essentially minimize the number of failures and thus lower a number of maintenances after failure, thereby the consumptions of expensive spare parts (which tend to be unnecessarily drawn just during emergency breakdowns) will decrease
- Decrease of costs due to the lower overtime of maintenance personnel and cost reduction for outsourced maintenance services
- Reduction of overall maintenance cost over a long time period [4]
- Based on collecting data related to failures and their analysis – repeated failures are significantly reduced
- A number of non-conforming products will decrease and thus lower the cost of non-quality
- Furthermore, increasing the efficiency of maintenance results in reducing the environmental impacts of organization's activities and improves the safety and health at work which is the goal of the integrated quality management system.

Currently, it is already common that the maintenance management system is supported by the information system. Without support of MMIS, it is not possible to assess the effectiveness of maintenance (unreliable data) and besides that it is not possible to effectively manage and plan maintenance activities (proper data is missing). [1] If it is assumed that maintenance is supported by the MMIS, it is possible to get a number of indicators, which enable to quantify most of the benefits of MMIS - they are already certified in the industrial sector and there is no reason why the agriculture sector (despite certain specifics) should be an exception. Why such a system is not yet widespread in the agro-industrial complex?

2. Material and method

The general idea of MMIS is based on the principles of logistics management system, whose main objective is to plan, manage and control material and information flow in order to achieve the performance and economic goals. A substantial part of the logistics management system is an information system (IS), which goal is to capture, store, process and transmit data (actual and planned). It is beneficial not only for well organized documentation of the maintenance and other activities (it is a starting point, it is not desired goal), but also to save time in preparing and implementing maintenance activities, saving human resources, material and spare parts, quick reduction of weak points, reduction of nonconforming products, increasing of the reliability of production equipment, etc. Everything can be achieved on condition that the IS has been correctly selected (user-friendly, open to changes, but also stable, etc.), successfully implemented (the implementation phase is very important full support of top management and effective involvement of all interested personnel) and appropriately used. [3]

MMIS electronically supports arrangement of a number of documents (related to operation) – asset inventory, chronological record of machine maintenance, work orders, a stock numbers of spare parts, a chronological records of parameters, production plans and schedules, chronological records of activities of workers, etc. [6] The basic prerequisite for using MMIS is proper records about maintained objects (based on the "asset inventory"), which is connected to the planning and keeping track of maintenance activities with linkage to the consumption of inventory during maintenance. Following additional sources of information (maintenance personnel availability, equipment and tools disposability, the cost of maintenance, failure codes,

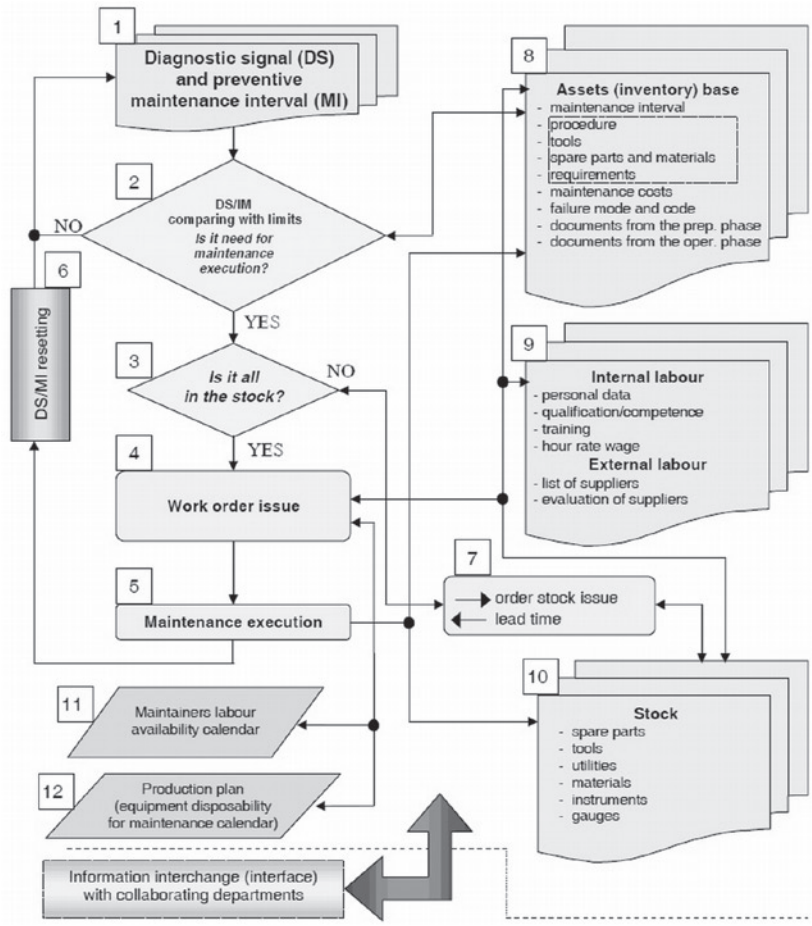


Fig. 1. Simplified MMIS flowchart of preventive maintenance

maintenance requirements, etc.) are recorded in long-planned and unplanned maintenance - their causes, labor, downtime, cost, etc. The maintenance data are commonly used for planning and various operational analysis of maintenance system - for example, to identify costs, labor intensity and duration of downtime of the selected object within selected period in order to determine the monthly maintenance costs for individual departments or production lines, to monitor the failure of machinery and eliminate frequent causes of failures, to summarize the maintenance personnel labor worked each year, etc. Knowing the theoretical bases of appropriate procedures (with recorded maintenance data), it is possible to obtain a number of other indicators that can significantly help in planning, assessment, and thus the continuous improvement of maintenance. Simplified MMIS diagram of preventive maintenance management in industries is shown in Figure 1.

Even agricultural company can use the same principle of MMIS, if the selection and implementation of MMIS will be done with respect of agricultural specifics versus industrial.

3. Results and discussion

Although it seems at first sight that maintenance systems in industrial and agricultural companies are very different, in fact they are not. Apparent obstacles using MMIS in agricultural companies exist in other sectors too, but these obstacles do not hinder using of MMIS - such as seasonal work exists in the sugar-houses, where the MMIS is used, as well weather conditions apply to construction companies, etc. The main causes of lower distribution of the maintenance management system in general and its computer support in agricultural companies are mainly these:

- In the Czech Republic there are relatively few large agricultural companies which could effectively use MMIS.
- Affordable, yet high-quality applications of MMIS appeared just recently on the market in the Republic
- Agricultural companies still do not appreciate the positive aspects of well-functioned maintenance system – similar situation was apparent in the industrial companies 15 years ago
- Lack of software and hardware facilities in the agricultural companies causes poor data communication among their divisions and also between agricultural companies and companies providing the maintenance outsourcing
- For complex machineries the maintenance is outsourced through dealers of specific brand and maintenance system is entirely managed by the dealer; for simple machineries maintenance after the failure is usually applied by own workers, often unskilled
- Lack of companies offering outsourced maintenance of agricultural machineries without a narrow link on a specific manufacturer or supplier of agricultural machinery.

However, it is possible to assume that in the up-coming years MMIS will be extended to larger agricultural companies. Some dealers of farm machinery have been working on applications that will ensure efficient flow of data within a communication channel dealer (repair service company) – agricultural company. These applications enclose evidence, planning and managing of maintenance activities, etc. Density of the service networks of suppliers of agricultural machinery is already sufficient, and it is still constantly growing. Agricultural companies (which use MMIS) and dealers of agricultural machinery are together connected via the corporate Web applications on a central server with data store and access to knowledge bases, spare part inventories, etc.

Effective maintenance systems should be based on collaboration and data link at least the level of largest suppliers of agricultural machineries in the Czech Republic with individual agricultural companies. Consolidated information system should share necessary data for maintenance management of particular agricultural companies, evidence of maintenance history and plans of preventive maintenance etc.

However, this solution is not an option because of concerns about misuse of enterprise data shared by cooperating companies. These concerns are often completely unnecessary, information technology allows access to data configuration for participating users, but in practice it is clear that sometimes claustrophobic fears about misuse of data will not be overcome in the near future.

A. Traditional in-house maintenance system

This model could be functional identically as in industry, where using the MMIS system of maintenance procedures is managed without relationship to the external environment and various maintenance subcontractors (outsourcing maintenance). Maintenance carried out by own workers are registered and MMIS planed essentially in the same way as outsourced maintenance. Maintenance system setting, the continuous monitoring and optimization of efficiency is controlled by the maintenance manager of agricultural company. This model is already running in several agricultural companies in the Czech Republic, but its disadvantage is the need to duplicate records of major machineries, for which maintenance is provided by external companies (usually dealers of the brand that have delivered machinery) who are also interested to record maintenance activities. By recording maintenance data on the same machinery in two different information systems often leads to errors due to human error, and quite often, the dealer has different data than the agricultural company. Agricultural companies are not usually equipped with devices for collecting diagnostic signals (which have dealers). Such a fact results in situation that agricultural companies do not have relevant and valuable information about the technical state of machinery - particularly for predicting the evolution of the technical condition of the machine and thus the need of future preventive maintenance.

B. Shared maintenance system

This system has the advantage that it eliminates the major insufficiencies of traditional in-house maintenance system. Agricultural company and dealer work together with the data collected on a single server, the MMIS set rights for individual users for access to input and view data. Unlike the previous model, the maintenance manager of the agricultural company could analyze the output of diagnostic device from dealer. On the other hand, the dealer should have a detailed overview of the planned maintenance which provides for the agricultural company, failure of machines, the need for spare parts, etc. The great advantage of this model is that if the planned maintenance in the MMIS is approved, both interested partners can prepare in advance. Affordable MMIS also includes failure monitoring module, so if a requirement is entered into MMIS by agricultural company, it immediately informs the dealer and the failure may be quickly removed. Even it is useful for both sides, described model is not spread in agricultural sector (minor share of this model in industry). This model is facing the above-mentioned reluctance to share data among partners.

C. Outsourcing of maintenance of agricultural machinery

Like in the industry, also in agriculture it is possible to expect the establishing of specialized companies for the maintenance of various agricultural machines. These companies could completely ensure the maintenance of machinery for agricultural companies by supply method. All data collected during maintenance of agricultural machinery would be owned and maintained by MMIS in outsourcing company. Using of MMIS would be in the interest of outsourcing company for easy planning of preventive maintenances and analysis capabilities of maintenances and also in order to increase its effectiveness. This model is typically used in the industry (chemical industry, heating plants, electric grids, etc.), but it has disadvantage, which is complicated communication in terms of dates of scheduled maintenance activities, and as well recovery of operational failures. There

Table 1. List of machinery and their average age and average downtime costs per hour

Machinery	Number of machinery	Average age in year 2010 (years)	Number of machinery		Average downtime costs per hour (CZK)
			In-house maintenance	Dealer's maintenance	
Tractors	28	12	13	12	900
Stubble ploughs	5	8	5	0	3 500
Ploughs	8	7	6	2	4 800
Machinery for soil cultivation before planting	22	14	22	0	3 400
Swing machinery	8	8	6	2	6 400
Machinery for fertilizing and crop protection	16	11	13	3	4 100
Harvesters + adapters	6	6	1	5	10 500
Forage machinery	21	7	16	5	3 850
Trucks	14	8	9	5	2 800
Sum, average	128	10	91	34	4 472

is a direct result of the absence of feedback dealers, and agricultural equipment manufacturers, for whom data of the maintenance of their machines during the life cycle is valuable informational basis for further development.

3.1. An example of usage of the MMIS to improve the quality of maintenance in agricultural company

The following example shows the benefits of implementation and subsequent maintenance improvement of agricultural machinery in a particular agricultural company. The company manages 3106 ha (7675 acres), has 112 employees, basic information about its machine park are listed in Table 1.

By the year 2003 (the maintenance was carried out without any system) the majority was corrective maintenance. Preventive maintenance including motor oil changes in mobile machines were made in a subjective decision of machine operator, maintenance documentation was not recorded. In 2004 a new management introduced system measures to improve the quality of maintenance of machinery for crop production. Initially the plan was to set up the proper planning of preventive maintenance and star to record maintenance activities. At the end of 2004, company acquired information system for maintenance management which began routinely used since 2005 but so far without sharing information with dealers of agricultural machines. Sharing information maintenance system with dealers for selected machines began to be applied from the beginning of 2009.

Already at the beginning of the implementation of MMIS in 2004, there were established criteria for assessing the effectiveness of maintenance. Furthermore, new indicators of efficiency maintenance were set up and their numerical values collected from 2005 to 2010 are listed in Table 2. Although there are indicators of maintenance efficiency commonly used in industrial plants, such as Overall Equipment Efficiency, however, these indicators can not be used in primary agricultural production. The newly proposed indicators are based on the requirements for an efficient maintenance system, which is expected to increase the dependability of machines and secondary factors namely availability (also durability and safety), maintenance supportability, while reducing costs of maintenance. When increasing the dependability can be expected less machine failures and thereby reduce downtime, maintenance costs of the failure and the cost of spare parts. When higher dependability is ensured, it can be expected less machine failures and thereby reduction of downtime, costs of corrective maintenance and costs of spare parts.

Additional commentary for the Table 2:

1. Direct annual cost of maintenance (obtained from MMIS) are calculated as a multiple of maintenance workload (intensity) and the maintenance man-hour rate, cost of used spare parts and maintenance material consumed is added to final amount of cost. Trend of direct annual costs of maintenance are shown in the Table 2 (row 1). Subdivisions of direct annual cost of maintenance are presented in rows 2-4.

Table 2. Evaluated indicators of maintenance effectiveness during the year 2005 to 2010

Item	Year					
	2005	2006	2007	2008	2009	2010
Direct annual maintenance cost (thousand CZK)	3 792	3 712	3 423	3 118	2 821	2 593
Cost of corrective maintenance (thousand CZK)	3 356	2 918	2 456	2 378	2 023	1 897
Cost of preventive maintenance (thousand CZK)	436	852	849	740	798	789
Cost of spare parts (thousand CZK)	2 156	2 351	1 860	1 721	1 693	1 456
Number of failures (-)	233	278	238	219	215	194
Downtimes caused by failure:						
Downtimes (hour)	552	499	481	452	406	302
Downtimes recounted on costs (thousand CZK)	2 469	2 232	2 151	2 021	1 816	1 351
Average cost of a corrective maintenance (thousand CZK)	14,40	10,50	10,32	10,86	9,41	9,78
Average length of a downtime caused by corrective maintenance (hour)	2,37	1,79	2,02	2,06	1,89	1,56
Recounted maintenance cost (thousand CZK)	6 261	5 944	5 574	5 139	4 637	3 944

2. The cost of corrective maintenance (obtained from MMIS) is the sum of all costs of maintenance in a given year (all maintenance performed because of failure of any machine).
3. The cost of preventive maintenance can be expressed as the sum of the costs of all preventive maintenance in a given year.
4. The cost of spare parts can be expressed as the sum of costs of all spare parts consumed during maintenance in a particular year.
5. Fifth row shows the number of failures per year, which were solved within maintenance activities. This indicator by itself is not conclusive, because it is not clear whether certain failure was a "big" (laborious, expensive) or "small". In this case, although the number of failures after 6 years has not fallen too much (from 233 to 194), but number of "big" failures was significantly reduced - this is evident in the development of the indicators listed in other rows of the table.
6. Downtimes due to failures are shown in other rows. They are given in hours and also recalculated to the cost of downtime.
7. Downtimes due to failures in hours - are expressed as the sum of downtime caused by corrective maintenance (obtained from MMIS).
8. Downtimes from the previous row are converted to costs by multiplying the average cost of an hour of downtime of the machine (these are shown in the last column of Table 1).
9. Average cost of corrective maintenance is obtained by dividing the total cost of corrective maintenance (row 2) by the number of failures per year (row 5).
10. Average downtime of corrective maintenance is obtained by dividing the total sum of downtime hours of corrective maintenance (row 2) by the number of failures per year (row 5).
11. Costs of downtimes of corrective maintenance are included in the annual cost of maintenance. Annual maintenance costs are calculated as the sum of direct maintenance costs (row 1) and the cost of downtime (row 8).

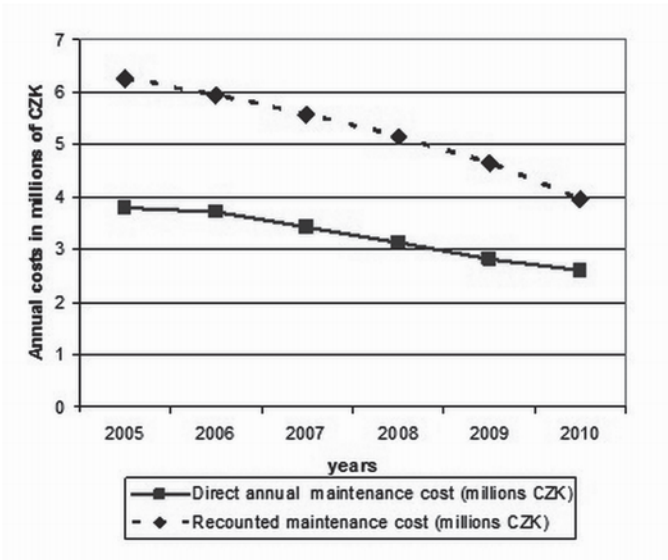


Fig. 2. Trend of direct and re-calculated annual maintenance costs

The trend of the indicators listed in the Table 2 shows that the implementation of a system with MMIS support can significantly reduce maintenance costs while increasing dependability of maintained machines. Especially the emphasis on preventive maintenance which is performed properly makes the biggest benefit of this system. Figure 2 shows that the direct costs for maintenance decreased by more than 1 million CZK per year (32%). Furthermore, re-calculated costs were even lowered by 2 million CZK (36%) during 6 years when all the data were tracked.

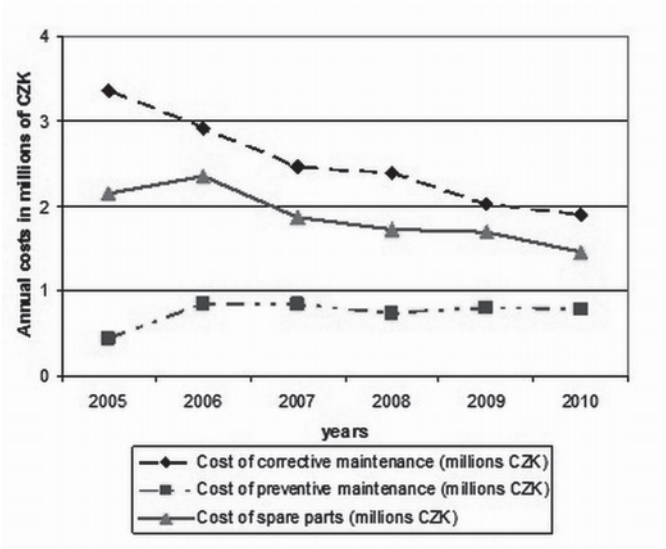


Fig. 3. Trend of subdivisions of direct annual maintenance costs

The trend of subdivisions of the direct maintenance costs (Fig. 3) is evident in each phase of improvements of the maintenance system. The cost of preventive maintenance after increasing of the intensity of preventive maintenance in 2006 abruptly doubled, the slight increase in 2009 is due to the implementation of preventive maintenance carried by dealer (year 2010 compared to year 2005: 172%). The increase of the cost of spare parts in 2006 is due to increased intensity of preventive maintenance and increase of quality of corrective maintenance, while their decline in 2010 was caused due to their lower purchase prices from dealer who performed certain maintenance (year 2010 compared to year 2005: 68%). The cost of corrective maintenance in 2006 and 2007 declined sharply due to the introduction of preventive maintenance in 2005, further decrease was more evident in 2009 after the dealer started to maintain more complex machinery (year 2010 compared to year 2005: 57%).

Improving the quality of the maintenance system with support of MMIS shared with the dealer is evident on the indicators in rows 9 and 10 in Table 2 - the average cost of corrective maintenance decreased by 32%, the average downtime for maintenance after a failure by 30%. It is obvious that although the number of failures decreased only by 17% in 2010, failures are less severe than in the year 2005.

In addition, MMIS and long-stored data about performed maintenance allow analysis of weak points on maintenance system – search of multiple failures and their elimination and possible change of spare parts suppliers with higher frequency failures, changes in schedules of post-season and pre-season maintenance based on the causes of failures, optimizing intervals of preventive maintenance etc.

5. Conclusion

Effective maintenance management in any sector must necessarily be supported by an information system that provides the necessary data for decision making based on real facts. MMIS are increasingly used to manage maintenance systems in the companies. Ten years ago use of MMIS was rather exceptional. MMIS was used scarcely only in industrial enterprises, are now it is almost the rule. Managers of these companies understand that properly implemented maintenance system is another source for the company to reduce costs, and that without the computer aided support the whole company can not achieve effective outcomes. MMIS are increasingly being used in larger agricultural companies in the Czech Republic, where using of

the traditional model of in-house maintenance system poses problems for outsourced maintenance of agricultural machinery dealers. In the following years, it is possible to assume that there will be closer cooperation between dealers and agricultural companies and expansion of a shared maintenance system of MMIS. Shared MMIS in the Czech Republic is still used in only a few agricultural companies. Nevertheless, as it is shown in the example, it is possible to reduce significantly costs while increasing dependability of agricultural machinery.

References

1. Eti M C, Ogaji S O T, Probert S D. Reducing the cost of preventive maintenance (PM) through adopting a proactive reliability-focused culture, Applied Energy 83 2006; p. 1235–1248.
2. Jurca V, Ales Z. Maintenance Management Efficiency Evaluation. Eksploatacja i Niezawodnosc – Maintenance and Reliability, 2007; 33(1); 13-19.
3. Kans M. An approach for determining the requirements of computerised maintenance management systems. COMPUTERS IN INDUSTRY. 2008; 59, p. 32-40.
4. Legat V, et al. Contribution to optimization of preventive replacement. Reliability Engineering and System Safety 51, Elsevier Science Limited, 1996; ISSN 0951-8320.
5. Michel H, Mufeed A. Improving industrial process safety & availability, Reliability Engineering 2008; p. 1021–1026.
6. Westerkamp T A. Maintaining maintenance - How smart managers plan and execute successful CMMS implementations. INDUSTRIAL ENGINEER. 2006; 38, p. 37-42.
7. Wiest W. Improvement of the availability through quality control and preventive maintenance of self propelled vehicles. In: Conference on Agricultural Engineering, VDI BERICHTE vol. 1449, Germany: 1998; p. 351-356.

Prof. Ing. Vladimir JURCA, C.Sc.
Czech University of Life Sciences Prague, Faculty of Engineering
Department for Quality and Dependability of Machines
Kamycka 129, 165 21 Prague – Suchdol, Czech Republic
E-mail: jurca@tf.czu.cz

Ing. Zdenek ALES, Ph.D.
Czech University of Life Sciences Prague, Faculty of Engineering
Department for Quality and Dependability of Machines
Kamycka 129, 165 21 Prague – Suchdol, Czech Republic
E-mail: ales@tf.czu.cz

Article citation info:
KRUCZYŃSKI S W. Maintenance of three way catalytic converter – thermal deactivation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 239-242.

Stanisław W. KRUCZYŃSKI

MAINTENANCE OF THREE WAY CATALYTIC CONVERTER – THERMAL DEACTIVATION

EKSPLOATACJA TRÓJFUNKCYJNYCH REAKTORÓW KATALITYCZNYCH – DEZAKTYWACJA TERMICZNA*

The present paper gives a description of deactivation processes of three way exhaust gas catalytic converters, with special consideration of thermal deactivation. Test results of new (or "fresh") and aged reactor during five-hours thermal test are presented. Conversion of carbon monoxide, hydrocarbons and nitric oxides is evaluated. Ignition temperatures of catalyst of individual noxious matters are determined, as well for new as for aged catalyst. Results of conversion measurements are correlated with results of physical and chemical tests of catalyst structure changes during ageing test.

Keywords: maintenance of catalytic converter, thermal deactivation, precious metal dispersion.

W pracy opisano procesy dezaktywacji trójfunkcyjnych reaktorów katalitycznych spalin ze szczególnym uwzględnieniem dezaktywacji termicznej. Przedstawiono wyniki badań rektora nowego i reaktora starzonego w pięciogodzinnym teście termicznym. Oceniono konwersję tlenku węgla, węglowodorów i tlenków azotu. Wyznaczono wartości temperatury zapłonu katalizatora poszczególnych substancji szkodliwych dla kata-lizatora nowego i starzonego. Wyniki pomiarów konwersji skorelowano z wynikami badań fizyko-chemicznych zmian struktury katalizatora podczas testu starzeniowego.

Słowa kluczowe: eksploatacja reaktorów katalitycznych, dezaktywacja termiczna, dyspersja metali szlachetnych.

1. Introduction

Modern catalytic converters are manufactured on the basis of ceramic or metal monolith, on which subsequent layers are applied, said layers performing different functions. They are composed of more than 5 different metal oxides and 2 - 3 kinds of precious metals. Reduction efficiency of carbon monoxide, hydrocarbons and nitric oxides emission for a fully functional catalytic system, running in steady state at appropriate temperatures and A-F mixture composition close to stoichiometric, exceeds 90%. Long-lasting use of the reactor causes its ageing and unavoidable deactivation process of catalytic layer. Changes of structure and chemical constitution of catalyst bed take place, with simultaneous deposition of layers of different chemical compounds, blocking the access to the catalyst.

Deactivation of exhaust gas catalytic reactor may result from different processes that may be divided into following groups [3]:

1. chemical processes, including adsorption of poison precursors and progressive poisoning, consisting of surface structure modifications and chemical blocking of active areas;
2. thermal processes, including changes of carrier and metal crystallites structure, sintering, oxidation and creation of precious metals alloys, and evaporation of metals;
3. mechanical processes, including generation of stresses as a result of thermal shocks and jolts causing friction and crushing of monoliths and carrier.

During normal operation of catalytic reactor, deactivation processes may proceed according to all above mentioned mechanisms. At actual level of manufacturing technique and operation of exhaust gas catalytic reactors, deactivation through mechanical processes takes little part to deactivation as a whole. Thermal deactivation together with chemical deactivation predominate [1].

2. Thermal deactivation tests of a converter

Tests were carried out using test stand for converter testing, constructed using electric tubular furnace. In the furnace a catalyst chamber, made of heat-resisting steel, was installed, into which exhaust gas from a Rover 1.4 engine are supplied via a heated gas path. Exhaust gas samples from upstream and downstream the catalyst are taken into exhaust gas analyzers. Exhaust gas temperatures are measured using thermocouples upstream and downstream the catalyst. Exhaust gas from downstream the converter were cooled, and after moisture condensing flowed through a set of rotameters.

Tests were performed using a Pd/Rh - Al2O3/CeO2 catalytic converter for small-engine car, made on the basis of metal monolith (number of ducts: 62 1/cm2) with precious metals amount of 1.46 g Pt/dm3 and 0.3 g Rh/dm3. For test purposes, the car catalyst was transformed into model test catalyst, having Φ = 30 mm (diameter) and l = 90 mm (length), then placed inside the tubular furnace. Ageing procedure was performed using exhaust gas [4] having stoichiometric composition (CO2 = 14.3%, CO = 0.56%, O2 = 0.65%, HC = 213 ppm, NOx = 2500 ppm), for 5 hours at 900°C with flow rate corresponding to relative volumetric flow velocity SV = 6000 h^{-1}. Before ageing procedure and after its termination a series of tests was performed in order to determine the correlation between the conversion of exhaust gas components and catalytic process temperature. Test results were completed with tests of selected characteristics of catalyst surface, including tests of porosity, palladium dispersion and X-ray structure tests. Tests are presented according to following methods and sequence.

1. Tests of porous structure of catalyst surface before and after ageing were carried out using ASAP 2010 apparatus. Catalyst fragments were examined (steel foil coated with oxide layer

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

on both sides with deposited precious metals). Test results refer to 1 g of catalyst mass.

2. X-ray structural tests were carried out sing Rigaku Denki X-ray diffractometer, in Bragg-Brentano reflective geometry, using CuK radiation and Ni filter. Tests were carried out on a catalyst in form of powder obtained from new and aged catalysts.
3. Precious metals dispersion was examined using hydrogen chemical adsorption method. Tests were carried out on powder previously obtained from new and aged catalysts. For final determination of crystallites dimensions and their degree of dispersion, one has to know the concentration of precious metals in catalysts being tested, said concentrations being determined using flame AAS method.
4. Conversion of gases was examined at constant gas flow with stoichiometric composition, controlled in closed loop using oxygen sensor, corresponding to relative volumetric flow velocity of exhaust gas SV = 12000 h^{-1} under engine operating conditions corresponding to n = 2500 RPM, and N_e = 25 kW effective power. Concentration of exhaust gas upstream and downstream the catalyst was measured depending on catalytic process temperature, and conversion of exhaust gas components was calculated, and then the ignition temperature T^{50} of catalyst was determined (temperature corresponding to 50% conversion of a given component).

3. Test results

Ad 1. Results of tests concerning determination of specific surface (BET model) and surface and volume of pores (BJH model) for new and aged catalyst are shown in table 1. The result of thermal deactivation of a catalyst is the loss of its active surface. A decrease of surface area (BET) and surface of pores (BJH) down to approx. 35% of initial value was found. The volume of pores decreased down to approx. 85% of initial value, whereas the average pore diameter increased from about 110 Å up to 240 Å. Such effects are due to the fact that pores having comparatively small diameter disappear as first.

Table 1. Results of tests concerning surface area and surface and volume of micro pores

Description	Unit	New catalyst	Aged catalyst
Surface area - BET model	m^2/g	22.3	7.74
Surface of pores - BJH model (adsorption)		23.3	8.9
Surface of pores - BJH model (desorption)		25.3	9.5
Volume of pores - BJH model (adsorption)	cm^3/g	0.068	0.056
Volume of pores - BJH model (desorption)		0.06	0.055
Average diameter of pores - BET model	Å	72.82	120.1
Average diameter of pores - BJH model (adsorption)		117.1	250.0
Average diameter of pores - BJH model (desorption)		107.6	232.0

Ad 2. Results of X-ray tests in form of Pd/Rh-Al_2O_3/CeO_2 catalyst diffraction pattern before ageing test (upper line) and after ageing test (lower line) are shown according to [1] in figure 1. In preparation Pd/Rh of new catalyst the mixture of γ-Al_2O_3¬ and CeO_2 was identified. Metal phases were not identified (except support of layers). In preparation of Pd/Rh catalyst after ageing test, γ-Al_2O_3 was still identified, as well as well-crystallized $CeAlO_3$ phase. Evaluation of crystallites size for that phase, obtained using Scherrer formula and half-width reflection, was equal from 340 to 400 Å for different reflections. In that preparation, reflection (111) from Pd phase (or rich in Pd alloy) was also observed. Small intensity of metal phases reflections being observed, and their superposition with dominating reflections from oxide phases make possible the only one statement: phases are identifiable. Measured thickness of layers deposited on the support was equal to 0,045 mm approx., whereas said layers after test showed certain irregularity of thickness (within 0,005 mm)

Ad 3. Test of metals dispersion preceded by temperature-programmed reduction (TPR) and temperature-programmed desorption (TPD) were carried out using hydrogen chemical adsorption method. These measurements were performed using impulse method. Hydrogen was injected in form of pulses into argon stream until surface saturation was obtained. Lack of hydrogen chemical adsorption manifested in "passing through" the sample of equal peaks, watched on recorder's screen. Hydrogen is chemically adsorbed in form of monolayer on metal surface. Chemical adsorption of hydrogen on metals of groups 8 − 10 (thus also on palladium and rhodium) takes place with dissociation, so it is assumed that one atom of chemically adsorbed hydrogen $H_{(ads)}$ falls on one surface atom of metal (H_{ads} / M_{pow} = 1, where M = Pd, Rh).

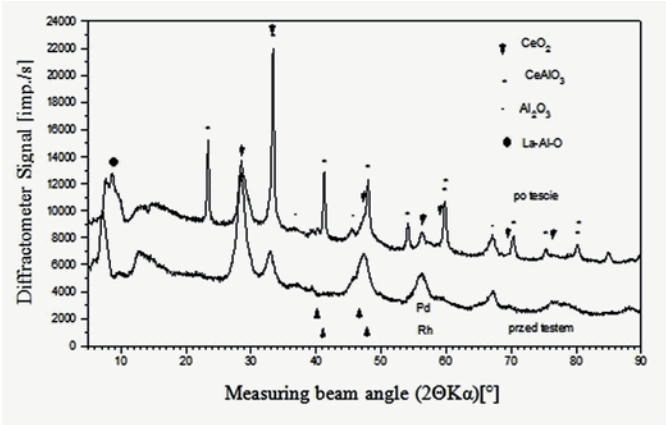


Fig.1. Diffraction pattern of Pd/Rh - Al2O3/CeO2 catalyst (powder) before ageing test (lower line) and after ageing test (upper line) [1].

Hypothetical, typical chemical adsorption image (figure 2) looks as follows: the 1st hydrogen pulse is partially adsorbed, whereas next pulses pass completely through the sample and may be taken as standard. The difference of areas under the first pulse, and average area under next pulses is the measure of adsorbed hydrogen amount. Knowing the amounts of hydrogen pulses injected into argon stream,

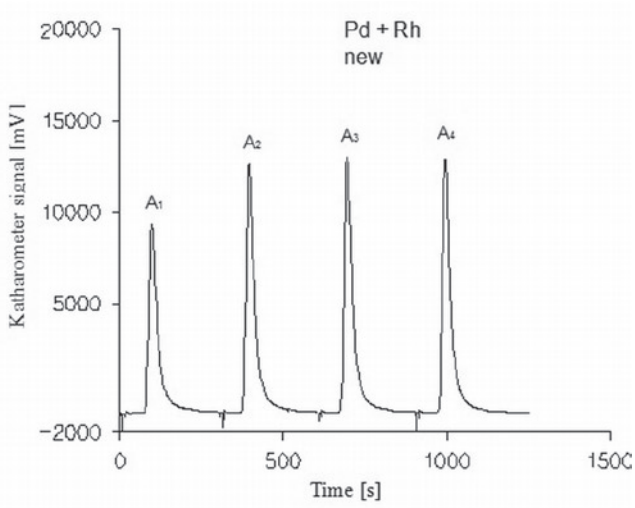


Fig. 2. Example of hydrogen chemical adsorption process on new Pd/Rh catalyst. A_1, A_2 , A_3 and A_4 mean areas under individual hydrogen pulses

Table 2. Results of metal dispersion measurements according to [2]

catalyst	catalyst mass / number of metal moles in sample	number of moles of adsorbed H_2		dispersion, D%		average crystal-lite size [Å]
		chemisorption after TPR	chemisorption after TPD	measurement after TPR	measurement after TPD	
Pd-Rh, new;	0.25157 g/ 13.735e-6	0.875e-6	1.180e-6	12.74	17.18	91
Pd-Rh, aged	0.31443 g/ 16.839e-6	0.591e-6	0.269e-6	7.02	3.19	268

one has calculated from ideal gas law the number of hydrogen moles adsorbed by catalyst sample. Then, assuming that hydrogen is adsorbed by surface metal atoms only, and assuming adsorption model as: 1 hydrogen atom per 1 metal atom, one has calculated average crystallite size and metal dispersion in the catalyst. Results of measurements according [4] are shown in table 2.

Results of metal crystallites size using hydrogen chemical adsorption method (see table 2.) have to be construed very carefully, as the may be distorted by different factors. Said factors are:

- metal crystallites may be enclosed by carrier layer, preventing hydrogen adsorption. Because of this one will calculate a smaller dispersion, thus larger metal crystallites,
- the metal in catalysts being used, even in initial state, may be "decorated" with other components, and additionally in aged catalyst, there may be carbon deposits on metal surface, coming from polymerization process of reaction being catalyzed. This situation may make difficult hydrogen adsorption, resulting in calculation of too large values for metal crystallites size.

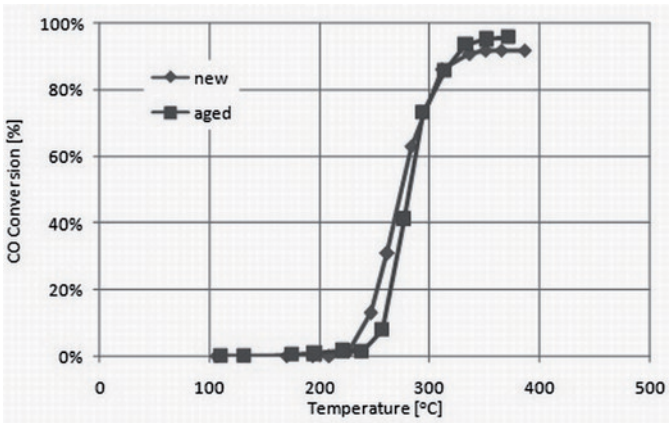


Fig. 3. Carbon monoxide conversion as function of temperature, for new and aged converter

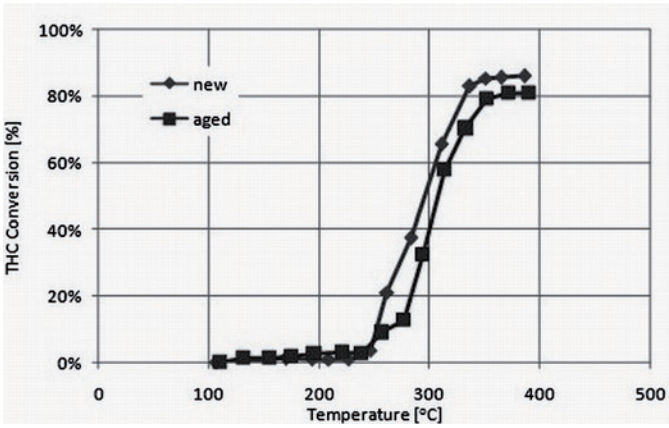


Fig. 4. Conversion of hydrocarbons as function of temperature, for new and aged converter

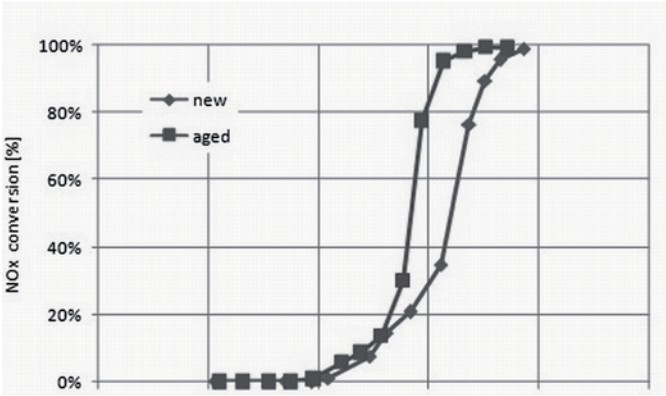


Fig. 5. Conversion of nitric oxides as function of temperature for new and aged converter

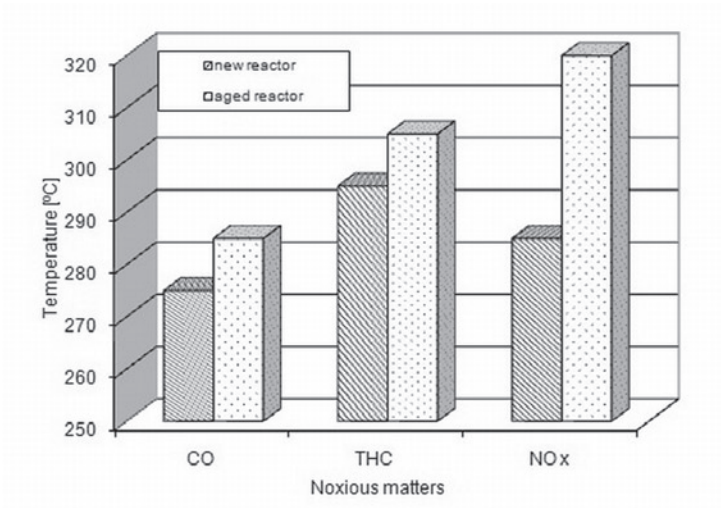


Fig. 6. Temperature comparison for 50% conversion of CO_2 THC and NO_X for new and aged converter

Ad 4. Results of measurements and calculations of carbon monoxide, hydrocarbons and nitric oxides conversion as function of temperature, for new catalyst and catalyst after ageing test are shown in fig. 3, 4 and 5.

Fig. 6 shows 50% conversion temperatures calculated for individual noxious matters, for new catalyst and catalyst after ageing test.

As a result of thermal deactivation process of a catalyst it was found, that there was a clear increase of 50% conversion temperature for each noxious matter, which was for carbon monoxide, hydrocarbons and nitric oxides 10°C, 10°C and 35°C respectively.

4. Conclusions

1. As a result of thermal deactivation test, the porosity of the catalyst decreased unfavorably, and in the same time small and medium pores were mainly destroyed. Total volume of pores, as well as total surface of pores were decreased.
2. In the catalyst after deactivation test a well-crystallized CeAlO3 phase was discovered, said phase could contribute to possible decrease of oxygen storage capacity in oxide layer of the carrier. Estimated sizes of crystallites in said phase, obtained using Scherrer formula and half-width reflection, were in the range from 340 to 400 Å for different reflections. In this preparation the appearance of reflection from Pd phase (or rich in Pd alloy) was also observed, a possible indication of Pd-Rh alloy being formed.
3. The thermal deactivation caused a clear increase of average size of precious metals crystallites. Despite reservations concerning the accuracy of this method it can be found, that the decrease of metals dispersion is undeniable and may be essential for catalyst activity.
4. It was found that the thermal deactivation of catalyst under test resulted in essential increase of 50 % conversion temperature of all noxious matters, from about 10°C up to about 30°C, which may cause essential increase of emission during homologation tests including cold engine start-up.

References

1. Birgerson H, Boutonnet M, Jaras S, Ericson L. Deactivation and regeneration of spent three way automotive exhaust gas catalysts. Topics in Catalysis, New York. 2004 30/31: 433-437.
2. Bonarowska M. Research dispersion catalysts Pd/Rh i Pt/Rh on Al_2O_3. Unpublished. Warsaw 1999. [In polish]
3. Kruczyński S. Three Way Catalytic Converter. ITE Radom 2004. [In polish]
4. Lassi U, Hietikko M, Rahkamaa-Tolonen K, Kallinen K, Savimaki A, Harkonen M, Laitinen R, Keisiki R. Deactivation correlations over Pd/Rh monoliths: role of gas phase composition. Topics in Catalysis, New York. 2004. 30/31: 457-462.
5. Pielaszek J. X-Ray structural analysis of catalysts Pd/Rh i Pt/Rh. ICHF PAN. Unpublished. Warsaw 1999. [In polish]

Prof. nzw. dr hab. inż. Stanisław W. KRUCZYŃSKI
Institute of Vehicles
Warsaw University of Technology
Ul. Narbutta 84, 02-524 Warsaw, Poland
e-mail: skruczyn@simr.pw.edu.pl

Article citation info:
ZARINCHANG A, FAGHIH N, ZARINCHANG J. An application of Genetic Algorithm toward solving the reliability problem of multiobjective series-parallel systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 243-248.

Amir ZARINCHANG
Nezameddin FAGHIH
Jafar ZARINCHANG

AN APPLICATION OF GENETIC ALGORITHM TOWARD SOLVING THE RELIABILITY PROBLEM OF MULTIOBJECTIVE SERIES-PARALLEL SYSTEMS

ZASTOSOWANIE ALGORYTMU GENETYCZNEGO DO ROZWIĄZYWANIA ZADAŃ NIEZAWODNOŚCIOWYCH DOTYCZĄCYCH WIELOKRYTERIALNYCH SYSTEMÓW SZEREGOWO-RÓWNOLEGŁYCH

Since developing an appropriate solution for reliability optimization problem with mathematical programming methods has been considered as difficult techniques, the heuristic approaches increasingly has been applied. Multiobjectve Genetic Algorithm (MGA) has been among heuristic methods that was developed to find solutions for series-parallel systems to obtain maximum reliability, and minimum cost and weight at the system level. These are very common problems in engineering design such as mechanical and electrical systems. It has been shown that the Multiobjectve Genetic Algorithm offers proper results to these problems while it respects to the several objective functions such as reliability, cost and weight. This paper presents the combination of probabilistic search, and one of the decision making methods called Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). The Multiobjectve Genetic Algorithm, allows us to achieve a proper design solution while it saves a considerable time compared with some other approaches. At the same time as the reliability, cost and weight were chosen as objective functions, the results obtained by this method showed an overall improvement in comparison to the existing GA method considering cost and weight as constraints.

Keywords: Multiobjective Genetic Algorithm, Reliability Optimization, Redundancy Apportionment, Series-Parallel Systems, TOPSIS Method.

Ponieważ znalezienie odpowiedniego rozwiązania zadania optymalizacji niezawodnościowej przy wykorzystaniu metod programowania matematycznego uznaje się za trudne, coraz częściej stosuje się do tego celu metody heurystyczne. Algorytm genetyczny do optymalizacji wielokryterialnej (Multiobjective Genetic Algorithm, MGA) jest jedną z metod heurystycznych, stworzoną w celu znajdowania rozwiązań dla systemów szeregowo-równoległych, pozwalającą na uzyskanie maksymalnej niezawodności oraz minimalnych kosztów i ciężaru na poziomie systemu. Zadania takie występują powszechnie w dziedzinie projektowania i konstrukcji systemów mechanicznych i elektrycznych. Wykazano, że MGA pozwala uzyskać odpowiednie rozwiązania tego typu zadań uwzględniając przy tym funkcje celu, takie jak niezawodność, koszty i ciężar. W niniejszej pracy przedstawiono połączenie metody wyszukiwania probabilistycznego oraz jednej z metod rozwiązywania problemów decyzyjnych o nazwie TOPSIS (Technique for Order Preference by Similarity to Ideal Solution). MGA pozwala uzyskać odpowiednie rozwiązania konstrukcyjne dając przy tym znaczną oszczędność czasu w porównaniu z niektórymi innymi metodami. Jednocześnie potraktowanie kosztów i ciężaru jako funkcji celu daje lepsze wyniki w porównaniu do metody wykorzystującej algorytm genetyczny, w której koszty i ciężar rozpatrywane są jako ograniczenia.

Słowa kluczowe: *Algorytm genetyczny do optymalizacji wielokryterialnej, optymalizacja niezawodności, podział nadmiarowości, systemy szeregowo-równoległe, metoda TOPSIS.*

1. Introduction

Reliability is an important factor in different kinds of electrical and mechanical systems. In many practical system designs, the overall system is divided to certain series parts called subsystems, according to the function requirements of the system (Fig. 1). For each subsystem, there are several parallel positions that can be filled with different component types available with varying reliability, costs, weight, volume and other characteristics [13]. Reliability of a system is calculated after computing the reliability of each subsystem. For optimizing system reliability, the following approach can be applied: (a) using more reliable components or applying better technical and organizational actions such as condition monitoring systems [19], (b) using redundant configuration with active or stand-by components in parallel, or (c) a combination of (a) and (b) [9], [13]. A well-known and complex reliability optimization problem is the Redundancy Apportionment Problem (RAP) for series-parallel systems which can be identified as the selection of the optimal combination of component type and redundancy level for each subsystem in order to meet various objectives on the overall system [13]. Conflicting objectives, such as minimizing the system cost and system weight or volume, while simultaneously maximizing the system reliability, make this problem complex. The RAP has proven to be NP-hard [2], [13]. Some approaches have been tried out to solve this problem [13], [14], [20]. The increase of the size and number of constrains affects computational difficulty exponentially, therefore, the majority of previous methods restricted the problem in a way. For instance, the number of components that could be chosen for a subsystem should not be more than one component. It means for providing redundancy only the same type can be used. Integer programming, dynamic programming, mixed integer and nonlinear programming, and multiobjective approaches could not find a solution if they did not apply restrictions. A method that could overcome these obstacles was Genetic Algorithm. Coit and Smith [3], [4], [13] solved the RAP problems by using Genetic Algorithm (GARAP). The

results of GARAP have been much better. Through many generations GA collects best solutions with an improving strategy and this method even results for large-scale problems [13]. In the existing GARAP cost and weight are considered as constraints, however, in this paper cost and weight are considered as objective functions.

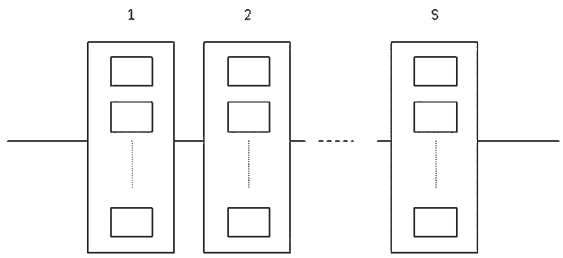


Fig.1. A Series-Parallel System

2. Notations, acronyms and assumptions

A. Notations

R_s	reliability of system
C_s	cost of system
W_s	weight of system
s	number of subsystems
m_i	number of available component choices for subsystem i (i = 1,...,s)
r_{ij}	reliability of j th component for subsystem i
c_{ij}	cost of j th component for subsystem i
w_{ij}	weight of j th component for subsystem i
n_i	total number of components used in subsystem i
$n_{\max}$	maximum number of components in parallel (user specified)
$R_i(x_i)$	reliability of subsystem i
$C_i(x_i)$	total cost of subsystem i
$W_i(x_i)$	total weight of subsystem i
v_j	vector encoding of solution j
λ_r	importance of system reliability (weight of reliability in TOPSIS method)
λ_w	importance of system weight (weight of weight in TOPSIS method)
λ_c	importance of system cost (weight of cost in TOPSIS method)
p	population size
f(v_j)	fitness for j th member of the population

B. Acronyms

GA(s)	Genetic Algorithm(s)
MGA	Multi-Objective Genetic Algorithm
RAP	Redundancy Apportionment Problem
OIMP	Overall Improvement

C. Assumptions

1. The reliability of each component is known.
2. Components are not repairable.
3. Failure of each component do not depends on other components.

3. Problem formulation

For modeling an RAP for the series-parallel systems, three objective functions must be optimized, that is, to maximize reliability (1) and minimize cost (3) and weight (4). At least on component in parallel is specified for each subsystem to operate. This problem can be easily expanded by adding more objective functions and be solved by the GA - the feature that previous formulations did not have [3] [10].

$$Max\ R_s = \Pi_{i=1}^{S} R_i(x) \tag{1}$$

$$R_i = (1 - \Pi_{j=1}^{n_{\max}} (1 - r_{ij})) \tag{2}$$

$$Min\ C_s = \sum_{i=1}^{S} \sum_{j=1}^{n_{\max}} c_{ij} \tag{3}$$

$$Min\ W_s = \sum_{i=1}^{S} \sum_{j=1}^{n_{\max}} w_{ij} \tag{4}$$

The total number of unique system configurations is given by the following equation [7], [13]:

$$N = \Pi_{i=1}^{S} \left[\binom{m_i + n_{\max}}{m_i} - \binom{m_i + 1 - 1}{m_i} \right] \tag{5}$$

4. Genetic algorithm implementation

GA is a stochastic global search method that mimics the process of natural biological evolution [6], GA operates on a population of potential solutions applying the principle of survival of the fittest to generate (hopefully) better and better approximations to a solution [21]. The steps of GA are [3], [15]:

1. Encoding of solutions
2. Initialing a population of chromosomes
3. Selecting parents for breeding
4. Creating new chromosomes by transferring best strings and crossover operator; applying mutation as the parents mate
5. If the termination criterion is met, stop and return the best chromosome; If not go to 3

Crossover and mutation operators play an important role in the GA. The rate of convergence is specified by the effectiveness of the crossover operator at the same time as the mutation operator restrains the algorithm not to stop in a local optimal and transferring best strings help algorithm to keep the best solutions in the next generations. The number of subsystems, maximum number of components in parallel, transferring and mutation rate is tunable parameters but constant in specific experiment [3], [4].

A. Solution Encoding

Although traditional GA encoded solution using a binary string [3], [11] for combinatorial optimization it is preferred and more efficient to encode solution with integer values. In this approach the second encoding method has been selected. Each subsystem includes n_i parts in parallel ($1 \le n_i \le n_{\max}$) in order to form a possible solution. The n_i parts can be selected from m_i components that are available for i th subsystem. Components are sorted according to their reliabilities from 0 to

($m_i - 1$), that is, the most reliable component is shown by 0 and the least reliable is shown by, $m_i - 1$. Chromosome is a vector with ($n_{\max}$ * s) positions. It means for each of s subsystems there are $n_{\max}$ positions that can be filled with an integer number. An index −1 is assigned to the empty position where there is no component (i.e., $n_i < n_{\max}$). At last all the subsystem representations are placed next to each other and form a vector that shows a chromosome [3]. As an example, consider a system with $s = 3$, $m_1 = 4$, $m_2 = 3$, $m_3 = 4$ and $n_{\max}$ predetermined to be 5. The following example:

$$v_j = (3\ 3\ 3\ -1\ -1 \mid 0\ 0\ 1\ -1\ -1 \mid 2\ 1\ -1\ -1\ -1)$$

represents a prospective solution with three of the fourth most reliable components for the first subsystem; two of the most reliable component and one of the second most reliable component used in parallel for the second subsystem; and the one of the third most reliable and one of the second most reliable component used in parallel for the third subsystem.

B. *Initial Population*

When user determined population size (p) and the number of subsystems (s), the initial population could be generated. For each gene (position) an integer between −1 and (m_i−1) was generated randomly and uniformly (with replacement) according to which subsystem this gene belonged to (to determine m_i). Previous article [5] indicated that a population size of 40 converged quickly and produced good solutions. In general, the minimum effective population size would grow with problem size [3], [11].

C. *Objective Function*

By generating initial population, p vectors were produced that each one demonstrates a system design. Reliability, weight and cost of generated solutions are calculated according to (1), (2), (3) and (4), therefore, a (p*3) primary matrix can be formed as illustrated in (fig.2).

(R_{sj}, W_{sj}, C_{sj}) Show reliability, weight and cost of jth design.

$$\begin{bmatrix} R_{s1} & W_{s1} & C_{s1} \\ R_{s2} & W_{s2} & C_{s2} \\ . & . & . \\ . & . & . \\ R_{sp} & W_{sp} & C_{sp} \end{bmatrix}$$

Fig. 2.Primary matrix

As primary matrix shows, there are three attributes for each alternative, so, they are not easily ranked. Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is applied to overcome ranking problem. Hwang and Yoon [12] developed the TOPSIS technique based on the concept that "the best alternative should have the shortest distance from the positive–ideal solution and the longest distance from the negative-ideal solution" and the ideal solution is the collection of ideal scores (or ratings) in all attributes considered [1]. Therefore, to use TOPSIS the following steps should be passed:

Step 1) Normalizing the Primary Matrix (Fig.3)

Through this step, three attribute dimensions transform to non-dimensional attributes which allow comparison across criteria.

$$R'_{sj} = \left[R_{sj} \Big/ \left(\sum_{j=1}^{p} R_{sj}^2 \right)^{\frac{1}{2}} \right] \tag{6}$$

$$W'_{sj} = \left[W_{sj} \Big/ \left(\sum_{j=1}^{p} W_{sj}^2 \right)^{\frac{1}{2}} \right] \tag{7}$$

$$C'_{sj} = \left[C_{sj} \Big/ \left(\sum_{j=1}^{p} C_{sj}^2 \right)^{\frac{1}{2}} \right] \tag{8}$$

$$\begin{bmatrix} R'_{s1} & W'_{s1} & C'_{s1} \\ R'_{s2} & W'_{s2} & C'_{s2} \\ . & . & . \\ . & . & . \\ R'_{sp} & W'_{sp} & C'_{sp} \end{bmatrix}$$

Fig.3. Normalized matrix

Step 2) Calculating Weighted Normalized Matrix (Fig.4)

Since attributes have different importance, a set of weights (λ_r, λ_w, λ_c) are assigned to them depends on how important the attributes are for users. It should be considered that sum of weights is equal to one.

($\lambda_r + \lambda_w + \lambda_c = 1$)

$$R''_{sj} = \left(\lambda_r \times R'_{sj} \right) \tag{9}$$

$$W''_{sj} = \left(\lambda_w \times W'_{sj} \right) \tag{10}$$

$$C''_{sj} = \left(\lambda_c \times C'_{sj} \right) \tag{11}$$

$$\begin{bmatrix} R''_{s1} & W''_{s1} & C''_{s1} \\ R''_{s2} & W''_{s2} & C''_{s2} \\ . & . & . \\ . & . & . \\ R''_{sp} & W''_{sp} & C''_{sp} \end{bmatrix}$$

Fig. 4. Weighted Normalized matrix

Step 3) Identifying Positive-Ideal and Negative-Ideal Solutions

$$I^+ = \left\{ \max R''_{sj}, \min W''_{sj}, \min C''_{sj} \right\} = \left\{ R_s^+, W_s^+, C_s^+ \right\}$$

I^+ is a set of the best values for attributes among all alternatives. Best values for 2nd and 3rd columns are their minimum values since the least is the best.

$$I^- = \left\{ \min R''_{sj}, \max W''_{sj}, \max C''_{sj} \right\} = \left\{ R_s^-, W_s^-, C_s^- \right\}$$

I^- is a set of the worst values for attributes among all alternatives.

Step 4) Calculating Separation Measures

The distances d_j^+ and d_j^- to I^+ and I^- for all solutions are correspondingly computed according to (13) and (14):

$$d_j^+ = \left[\left(R''_{sj} - R_s^+\right)^2 + \left(W''_{sj} - W_s^+\right)^2 + \left(C''_{sj} - C_s^+\right)^2 \right]^{1/2} \quad (13)$$

$$d_j^- = \left[\left(R''_{sj} - R_s^-\right)^2 + \left(W''_{sj} - W_s^-\right)^2 + \left(C''_{sj} - C_s^-\right)^2 \right]^{1/2} \quad (14)$$

Step 5) Calculating the Relative Closeness to the Ideal Solution

The TOPSIS technique defines a "Similarity index" (or relative closeness) by combining the closeness to the positive-ideal solution and the remoteness of the negative-ideal solution. A relative distance D_j^+ comprised between [0,1] is assigned to each alternative which is the GA fitness function f(v_j) in this method. (The more relative closeness to the Ideal Solution means the solution is better)

$$D_j^+ = d_j^- / (d_j^+ + d_j^-) \quad (15)$$

Figure 5 [16] illustrates how TOPSIS works. Given an alternative like a_j, the distances d_j^+ and d_j^- to I^+ and I^- correspondingly are computed, subsequently a relative distance D_j^+ comprised between [0,1] is assigned to each alternative.

D. *Transferring Best Strings*

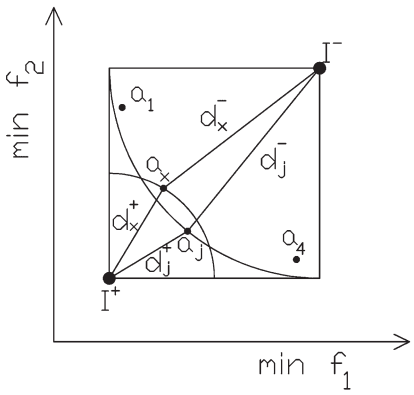


Fig. 5. TOPSIS distances [16]

Transferring best strings helped GA not to lose the best solutions in iterations. Transferring rate (0<TR<1) predetermined by user, therefore, in each generation (p * TR) chromosomes were kept without any changes according to their objective function values.

E. *Selection of Parents and Crossover Breeding Operator*

Crossover is the most important search operator in Genetic Algorithm, as it is a recombination operator which constructs two offspring out of genetic information encoded in two selected parents [21]. Roulette wheel [15] mechanism is employed by many selection techniques to select individuals on the basis of some measure of their performance probabilistically. The basic roulette wheel selection method is stochastic sampling with replacement (SSR) [15]. In each generation a string has an objective or fitness function value and sum of these values formed total fitness (16). v_j could be selected as a parent with the probability of $\frac{f(v_j)}{totalfit}$ based on roulette wheel method.

$$\text{Totalfit} = \sum_{j=1}^{p} f(v_j) \quad (16)$$

This crossover operator is a variation of the Single-Point operator which has been shown [22] to be superior to the Uniform crossover strategies for analyzing these problems. It also prevents creation of infeasible offspring during evolutions. In Single-Point operator same cutting point is selected randomly in both parents and then all data before the point is swapped between the parents. Two generated chromosomes are offspring.

F. *Mutation Operator*

In simple Genetic Algorithm [21], mutation is the less used but not the less important operator in comparison to crossover. Each gene in an offspring which has been created by crossover can be mutated with respect to the mutation probability (mutation rate) which is predetermined by user according to the size of the problem (length of the chromosome) [21]. A mutated component was changed to (its index + 1) and the last component (m_i) was changed to a position where an additional component was not used (−1).

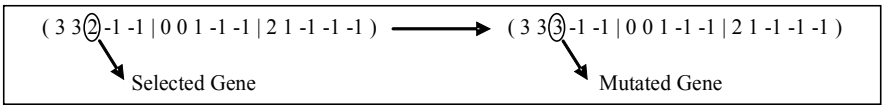


Fig. 6. Mutation Operator

G. *Evolution*

A survival of the fittest strategy is applied [3]. (p * TR) of the best solutions were copied to the next generation without any changes (transferring best strings) and the rest of new population that created by crossover operator were [p * (1 − TR)]. Mutation had its effect while producing new offspring with crossover operator. The best strings were never mutated because the best solutions should not be altered via GA. Since the GA is a stochastic search method, it is difficult to certainly find termination criteria. A common practice is to terminate the GA after a preselected number of generations in spite of having reached optimal solution much earlier [3], [15].

5. Test problem and results

A. *Test Problem*

The GA was used to analyze a different problem with very good results so it was implemented on the problem of the Fyffe, Hines and

Table 1. Best solution

Best Solution Vector	R_{sb}	W_{sb}	C_{sb}
002-1-1-10-10-1232-1-1-10-11011-1-1-13-1-1-13-1-11-1110-10-100-1-1-1-112-11-1-10-10101-100-1-10-1-1-1-13-1	0.9572	150	94

Table 2. Coit and Smith results

No.	Cost (C_{sa})	Weight (W_{sa})	Reliability (R_{sa})	OIMP%	No.	Cost (C_{sa})	Weight (W_{sa})	Reliability (R_{sa})	OIMP%
1	130	191	0.95675	0.95	15	123	174	0.97435	0.60
2	129	190	0.95603	0.90	19	122	173	0.97362	0.54
3	130	159	0.95556	0.94	20	120	172	0.97266	0.43
4	130	155	0.95503	0.93	21	121	171	0.97156	0.50
5	129	157	0.95429	0.59	22	120	170	0.97076	0.46
6	125	156	0.95362	0.53	23	120	169	0.96922	0.52
7	130	155	0.95311	0.93	24	119	165	0.96513	0.45
5	125	154	0.95239	0.52	25	115	167	0.96634	0.50
9	130	153	0.9519	0.92	26	116	166	0.96504	0.40
10	126	152	0.95102	0.71	27	117	165	0.96371	0.52
11	125	151	0.95006	0.55	25	115	164	0.96242	0.42
12	129	150	0.97942	0.90	29	114	163	0.96064	0.42
13	125	179	0.97906	0.64	30	114	162	0.95912	0.45
14	127	175	0.9751	0.75	31	113	161	0.95503	0.43
15	125	177	0.97715	0.65	32	114	160	0.95567	0.63
16	124	176	0.97642	0.62	33	110	159	0.95432	0.35
17	122	175	0.97552	0.51					

Lee problem [3], [8]. Five trials were performed for 1200 generation because the nature of GAs is stochastic, the final solution was found among the each and the best solution of five trials [3]. The results by the Multiobjective GA were used to compare its performance to existing GA. The size of the search space in this problem is greater than 7.6 x $10^{33}$from (5) [13].

B. Results

The best solution is showed (where (λ_r λ_w , λ_c) = (0.8, 0.1, 0.1)) in Table 1 and it is compared with results of Coit and Smith [4] in Table 2 which only considered reliability as an objective.

The overall improvement (OIMP %) is calculated according to (17) :

$$OIMP \quad (\%) = \begin{bmatrix} (R_{sb} - R_{sa}) / R_{sa} * \lambda_r + (W_{sa} - W_{sb}) / \\ W_{sa} * \lambda_w + (C_{sa} - C_{sb}) / C_{sa} * \lambda_c \end{bmatrix} * 100 \qquad (17)$$

6. Conclusion

GA has been demonstrated as a useful approach for solving Redundancy Apportionment Problem. In this paper we introduced MGA to provide a solution for systems that had to consider more than one objective, which since then it had not been reported that this problem was solved under this Multiobjective Genetic Algorithm formulation. Although the reliability of this result was not as proper as the previous formulation, the overall result showed a relative improvement. This algorithm did not use a complex decision making technique or local search to improve solutions but it seems that considering these features provides opportunities to have better results and more effective and efficient MGA.

References

1. Azar F. S. Multiattribute Decision-Making, use of three scoring Methods to compare the performance of imaging techniques for breast cancer detection, Technical Report, MS-BE-00-01MS-CIS-00-10, 2000.
2. Chen M. S. On the computational complexity of reliability redundancy allocation in a series system. Operation Research Letters 1992; 11:309–315.
3. Coit D. W, Smith A. E. Reliability optimization of series-parallel systems using a genetic algorithm. IEEE Transaction on Reliability 1996; 45(2): 254–260.

4. Coit D. W, Smith A. E. Penalty guided genetic search for reliability design optimization, Computer and Industrial Engineering 1996; 30(4):895–904.
5. Coit D. W, Smith A. E. Use of a genetic algorithm to optimize a combinatorial reliability design problem, Proceedings of the 3rd International Engineering Research Conference 1994; 467-472.
6. Ebson H. A genetic algorithm for macro cell placement, Proceeding of the European Design Automation Conference 1992; 52-57.
7. Feller W. An introduction to probability theory, New York, Wiley 1965.
8. Fyffe D. E, Hines W.W, Lee N.K. System reliability allocation and a computational algorithm, IEEE Transactions on Reliability 1965; vol R-17, 64-69.
9. Gen M, Cheng M. Genetic algorithms and engineering design, John Wiley & Sons, New York, 1997.
10. Glaß M, Lukasiewycz M, Haubelt C, Teich J. Lifetime Reliability Optimization for Embedded Systems: A System-Level Approach, Proceedings of International Workshop on Reliability Aware System Design and Test 2010; 17-22.
11. Goldberg D. E. Genetic Algorithms in Search, Optimization & Machine Learning, Addison Wesley, 1959.
12. Hwang C. L, Yoon K. Multiple Attribute Decision Making: Methods and applications, A state of the art survey, Springer-Verlag, Berlin, 1951.
13. Jian-Hua Z, Zhaoheng L, My-Thien D. Reliability optimization using multiobjective ant colony system approaches, Reliability Engineering and System Safety 2007; 92(1): 109-120.
14. Kuo W, Rajendra Prasad V. An annotated overview of system reliability optimization, IEEE Transaction on Reliability 2000; 49(2): 176-187.
15. Lawrence D. Handbook of genetic algorithm, Van Nostrand Reinhold, 1991.
16. Méndez M, Galván B, Salazar D, Greiner D, Multiple-Objective Genetic Algorithm using the Multiple Criteria Decision Making Method TOPSIS ,7th International Conference on Multi-Objective Programming and Goal Programming, 2006; 145-150.
17. Syswerda G. Uniform crossover in Genetic Algorithms, Proceedings of the 3rd International Conference on Genetic Algorithms 1959; 2-9.
18. Tate D. M, Smith A. E. A genetic approach to the quadratic assignment problem , Computers and Operations Research 1994; vol. 22, 73-53.
19. Tian Z, Levitin G, Zuo M. J. A joint reliability redundancy optimization approach for multi-state series–parallel systems, Reliability Engineering and System Safety 2009; 1568–1576.
20. Tillman F. A, Hwang C. L, Kuo W, Optimization techniques for system reliability with redundancy a review, IEEE Transaction on Reliability 1977; 26(3):145–55.
21. Zalzala A. M. S, Fleming P. J. Genetic algorithms in engineering systems, The Institution of Electrical Engineers, London, 1997.
22. Zarinchang A. Applications of Genetic Algorithm for Solving the Reliability Problem of Multi-Objective Series-Parallel Systems, MS Thesis, 2008.

Mr. Amir ZARINCHANG, Lecturer
Department of Industrial Engineering
Shiraz Branch, Islamic Azad University
Shiraz, Iran
amir_zarinchang@yahoo.com

Prof. Nezameddin FAGHIH, Professor
Department of Management
Shiraz University
Shiraz, Iran
faghihnezam@ut.ac.it

Dr Jafar ZARINCHANG , Associate Professor
Department of Mechanical Engineering
Shiraz University
Shiraz, Iran
jmzarin@yahoo.com

Article citation info:
NIEWCZAS A M, PIENIAK D, OGRODNIK P. Reliability analysis of strength of dental composites subjected to different photopolymerization procedures. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 249-255.

Agata M. NIEWCZAS
Daniel PIENIAK
Paweł OGRODNIK

RELIABILITY ANALYSIS OF STRENGTH OF DENTAL COMPOSITES SUBJECTED TO DIFFERENT PHOTOPOLYMERIZATION PROCEDURES

ANALIZA NIEZAWODNOŚCIOWA WYTRZYMAŁOŚCI KOMPOZYTÓW STOMATOLOGICZNYCH PODDANYCH ZRÓŻNICOWANYM PROCEDUROM FOTOPOLIMERYZACJI*

Abstract: The aim of this study was evaluation of chosen photopolymerization procedures on strength and reliability of dental composites based on siloranes and composites based on methacrylate compounds in 3-points bending test conditions. The following composites were tested: Filtek Siloran (FSi), Gradia Direct Anterior (GDA), Gradia Direct Posterior (GDP), Herculite XRV (H). Photopolymerization was conducted by means of two types of light: LED lamp and halogen lamp. Exposure times of 40 and 60 seconds were applied. For the strength studies a three-point bending test was used (TFS). Twenty rectangular beam-shaped samples (N=20) from each material were prepared for the studies. For each studied case an average value and standard deviation were determined. To assess significance of differences a variation analysis was performed. Then, the results from each specimen were approximated by two-parameter Weibull distribution. Distribution scale parameter was calculated (as a characteristic strength) and shape parameter (as a material reliability index). It has been demonstrated that in 3-point bending test conditions in case of silorane-based composite the type of lamp has no impact on the strength, however it can improve its reliability. In case of conventional methacrylate-based materials application of LED lamp instead of halogen lamp reduces material strength, but increases its reliability. Additionally, it has been shown that the extension of exposure time – in case of FSi material and halogen lamp, increases material strength, however it has no impact on reliability of the material.

***Keywords**: reliability, flexural strength, dental composites, photopolymerization.*

Celem badań była ocena wpływu wybranych procedur fotopolimeryzacji na wytrzymałość i niezawodność kompozytów stomatologicznych opartych na siloranach oraz kompozytów opartych na związkach metakrylanowych w warunkach testu na 3-punktowe zginanie. Badano kompozyty o nazwach handlowych: Filtek Siloran (FSi), Gradia Direct Anterior (GDA), Gradia Direct Posterior (GDP), Herculite XRV (H). Zastosowano fotopolimeryzację dwoma rodzajami światła: lampą diodową oraz lampą halogenową. Przyjęto czas naświetlania 40 sek. oraz 60 sek. Do badań wytrzymałości został zastosowany test na zginanie trójpunktowe (TFS). Przygotowano próbki do badań w formie belek prostopadłościennych o liczności N = 20 z każdego materiału. Wyznaczono wartość średnią i odchylenie standardowe dla każdego badanego wariantu. Do oceny istotności różnic przeprowadzono analizę wariancji. Następnie wyniki każdej próby aproksymowano dwuparametrowym rozkładem Weibull'a. Obliczono parametr skali rozkładu (jako wytrzymałość charakterystyczną) oraz parametr kształtu (jako wskaźnik niezawodności materiału). Wykazano, że w warunkach testu na 3-punktowe zginanie rodzaj lampy nie ma wpływu na wytrzymałość w przypadku kompozytu opartego na siloranach, natomiast umożliwia poprawę jego niezawodności. W przypadku konwencjonalnych materiałów opartych na metakrylanach zastosowanie lampy diodowej w miejsce lampy halogenowej obniża wytrzymałość materiału, jednak zwiększa jego niezawodność. Ponadto wykazano, że zwiększenie czasu naświetlania – w przypadku materiału FSi i lampy halogenowej zwiększa jego wytrzymałość, natomiast nie ma wpływu na niezawodność. W pozostałych przypadkach wytrzymałość na ogół pozostaje na stałym poziomie lecz zwiększa się niezawodność materiału.

***Słowa kluczowe**: niezawodność, wytrzymałość na zginanie, kompozyty stomatologiczne, fotopolimeryzacja.*

1. Introduction

Present-day laboratory studies of new biomaterials, such as dental composites, in many cases include not only studies of direct material characteristics but also prediction of preservation of these characteristics in operating conditions. Therefore, investigations of mechanical strength of composites are often extended by reliability analysis consisting in application of Weibull distribution as a failure probability distribution model, and estimation of the distribution parameters [20, 23, 27]. Reliability assessment is one of the components of complex risk analysis at clinical decision making by a dentist [11]. Weibull modulus (shape parameter) is adapted as a variation rate of material strength. A high Weibull modulus *m* indicates a potentially higher clinical reliability [4, 15]. The scale parameter of Weibull distribution specifies characteristic value of material strength, which corresponds to 63.2% of cases of failure of the studied material [3, 27]. Generally, a characteristic strength value (σ_0) depends on the material composition, photopolymerization and failure mechanisms [9].

Mechanical strength assessment of composites is usually per-

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

formed based on the bending tests. Three-point bending test (3PBT) is acknowledged by the International Organization for Standardization as a valid test for strength testing of dental polymer composites [9]. Three point bending can be compared to chewing process by molar and premolar teeth with regards to mechanics [8] – fig. 1.

Studies of flexural strength of dental composites can be used among the others for the assessment of photopolyimerization effectiveness, especially for evaluation of the effect of polymerization time and type of the lamp [1, 6, 10, 14, 20]. The most often applied lamps for polymerization in dental practice are LED and halogen lamps. The lamps using light-emitting diode are characterized by low energy consumption, moderate radiation intensiveness and very high durability.

Flexural strength test is a very important criterion of the clinical usability of composite materials. It is especially crucial in the context of dynamical development of dental composites, particularly introduction of silica fillers with particle sizes in a range of 0.1 nm to 100 nm, and siloranes as matrix material. Composites made of nano-particles with silorane matrix are characterized by a minimum polymerization shrinkage and fair mechanical properties (microhardness, flexural strength) [12,17,18] and satisfactory resistance to ageing and thermal fatigue [24,25]. However, in literature there is not much information available regarding new composites based on siloranes.

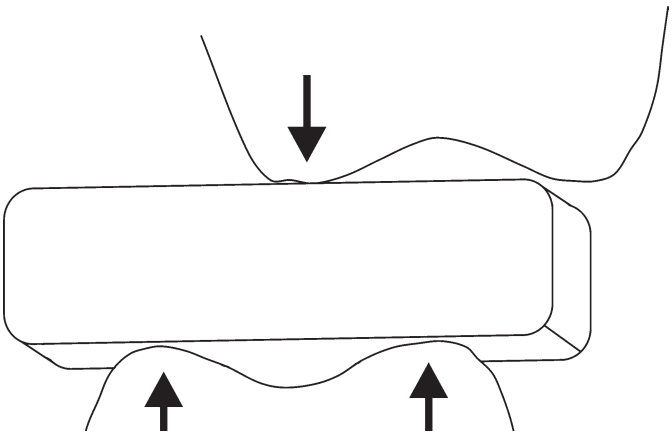


Fig. 1. Food crushing by three-point bending with molar and premolar teeth [8]

The aim of this study was to determine impact of photopolymerization technology, including exposure time and type of lamp, on the reliability of composites based on siloranes and standard methacrylate compounds, in laboratory flexural strength test conditions.

2. Materials and methods

The following commercial composites were studied: Filtek Silorane (3M ESPE) – FSi, Gradia Direct Anterior (GC) - GDA, Gradia Direct Posterior (GC) - GDP, Herculite XRV (Kerr) – H. Composites data can be found in Table 1.

Currently there are many dental light-cured composite materials available on the market. Most of them have a methacrylate matrix, which consists of few chemical compounds. A main group are monomers, such as for example Bis-GMA resin and its derivatives, urethane dimethacrylate and UDMA. The methacrylate matrix includes also comonomers, such as for example: TEGDMA and HEMA, which have a lower molar mass and reduce viscosity of basic resin.

In case of silorane based composites, silorane matrix consists of a hybrid, which is in half composed of silorane particles and oxiranes [12,17]. Such difference in matrix composition results in opening of silorane rings during polymerization, which causes their straightening and broadening, different from methacrylates, where monomers couple with each other by moving towards each other, resulting in significant volume reduction, and poses negative clinical effects due to polymerization shrinkage [7,12].

In case of both methacrylates and siloranes matrix the filler consists of silica based particles and fluoride aluminum silicate glass particles. Agent binding a resin with inorganic filler is most often organosilicon, vinyl and amine compounds.

Table 1. A list of studied composites

Material	Manufacturer	Type	Filler content (wt%)	Filler particles size
Filtek Silorane (FSi)	3M ESPE	Silorane	76	0,1 – 1 µm
Gradia Direct Ant (GDA)	GC	Microhybrid	73	average 0,85 µm
Gradia Direct Post. (GDP)	GC	Microhybrid	77	average 0,85 µm
Herculite XRV (H)	Kerr	Microhybrid	79	0,6 µm

The authors applied light-curing procedures by means of two type of lamps: LED lamp L.E. Demetron 1 (SDS/Kerr) and halogen lamp (Ivoclar Vivadent). Technical specification of the lamps is given in table 2.

Table 2. Specification of the lamps

Lamp	Type	Manufacturer	Wavelength [nm]	Power [mW/cm^2]
L.E.Demetron 1 (LED)	LED	SDS/Kerr	450÷470	200÷800
Astralis 7 (HAL)	Halogen	Ivoclar-Vivadent	400÷510	750

The applied exposure times were 40 and 60 seconds. Forty seconds is recommended by manufacturer of the materials. The authors introduced 60 seconds, based on the literature reports [5,13,21], which suggested that the extension of the exposure time improved polymerization effectiveness as a result of increase of the degree of conversion. The degree of conversion indicates a percentage of double bounding in polymer structure that underwent saturation (converted to single bond) [26]. Depending on the photopolymerization process conditions conversion degree is in a range between 65% and 80% [22,30].

2.1. Flexural strength studies

For flexural strength studies a three-point bending strength test was applied (TFS). Test samples (N = 20) were prepared in a shape of rectangular prism with dimensions of 2mm x 2mm x 25mm, according to PN-EN ISO 4049 standard. Tests were carried out at crosshead speed of 0.5 mm/min, with 20 mm distance between the supports. Supports radius and crosshead radius were 1 mm (Fig. 2).

Strength (σ) was calculated based on the following formula:

$$\sigma = \frac{3PL}{2bd^2}[MPa] \quad (1)$$

where:
P – loading during the test [N]
L – distance between the supports [mm]
b – specimen width [mm]
d – specimen thickness [mm]

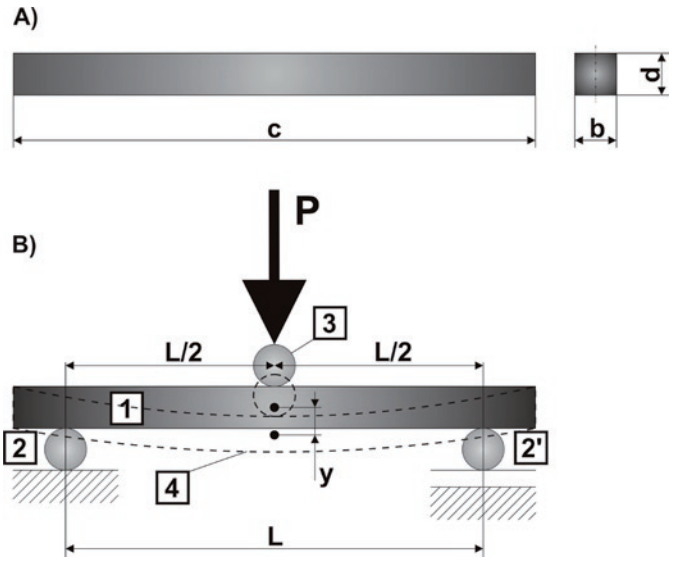


Fig.2. Scheme of the specimen (A) and test stand for the strength studies in 3-point bending test conditions - TFS (B): 1 – specimen; 2 – constant support; 2' – sliding support, 3 – loading crosshead; 4 – deflected beam; L – distance between supports; c, b, d, - specimen dimensions; y – beam deflection

2.2. Statistical analysis

Strength test results were divided into groups based on: composite type (FSi, GDA, GDP, H), type of lamp (HAL, LED) and exposure time (40 sec, 60 sec). In order to verify a significance of differences between sets of results a variation analysis of one variable was performed (ANOVA). *F – Snedecor* test was applied. Next, to evaluate a significance of direct differences between the sets assigned to particular photopolymerization procedures HSD Tukey's test was used.

For reliability analysis of the studies composites a two-parameter Weibull distribution was applied. Generally, a cumulative distribution function (P_f) of Weibull distribution (with positive parameters σ_0, m, and σ_u) is described by [16]:

$$P_f = 1 - \exp\left[-N\left(\frac{\sigma - \sigma_u}{\sigma_0}\right)^m\right] \tag{2}$$

where:
σ – failure load,
σ_0 – scale parameter,
m – shape parameter,
σ_u – location parameter,
e – constant (e = 2.71828...),
N – sample size,
P_f – probability of failure.

In case when the sample size N is constant in all studied sets (here specified by exposure time and type of lamp), N can be neglected in the calculations [2, 32].

If assuming location parameter value equal to zero $\sigma_u = 0$, Weibull distribution becomes two-dimensional. With these assumptions based on the equation (2) a formula for survival probability can be formulated, P_s:

$$P_s = 1 - P_f = 1 - \left(1 - \exp\left[-\left(\frac{\sigma}{\sigma_0}\right)^m\right]\right) \tag{3}$$

The above equation after finding a logarithm can be reduced to the linear form

$$y = ax + b \tag{4}$$

thus:

$$\ln\left[\ln\left(\frac{1}{P_s}\right)\right] = m\ln(\sigma) - m\ln(\sigma_0) \tag{5}$$

An intersection point of Y axis and approximating line depends on $-m\ln(s_0)$, while a slope parameter of approximating curve is a shape parameter of Weibull distribution m.

3. Test results

Descriptive statistics of flexural strength test results are given in table 3: sample size (N), average value, minimum and maximum value, standard deviation values, variation coefficient and Weibull modulus.

The highest average flexural strength value was obtained for FSi material after using HAL lamp and 60 seconds exposure time (119.5 MPa). Filtek Siloran (FSi) showed the highest strength in all photopolyimerization procedures (HAL 40, HAL 60, LED 40, LED 60).

Table 3. Descriptive statistic of flexural strength test results

Material	LCU	Time [s]	Descriptive statistic of RBC bending strength [MPa]					Co.Var. [%]
			N	Mean	Min	Max	St.Dev.	
FSi	HAL	40	20	108,15	85,10	126,00	9,00	8,32
		60	20	119,50	100,00	142,00	9,92	8,30
	LED	40	20	108,26	90,90	134,00	11,95	11,04
		60	20	107,76	82,60	143,00	13,15	12,20
GDA	HAL	40	20	77,77	63,90	86,60	5,52	7,10
		60	20	74,98	58,10	89,00	8,68	11,58
	LED	40	20	68,16	57,00	81,80	6,86	10,07
		60	20	67,48	41,80	83,30	9,35	13,86
GDP	HAL	40	20	84,15	76,60	91,00	3,74	4,45
		60	20	82,03	67,00	94,00	6,12	7,46
	LED	40	20	79,66	62,60	92,60	7,88	9,90
		60	20	76,08	56,10	95,70	11,97	15,73
H	HAL	40	20	107,23	69,90	125,00	13,26	12,37
		60	20	99,60	49,40	130,00	16,61	16,68
	LED	40	20	102,61	76,50	122,00	10,16	9,90
		60	20	102,05	79,90	124,00	12,09	11,84

The lowest strength in all groups was demonstrated for GDA material (67.48 MPa in LED 60 set).

The results of variance analysis obtained from F – Snedecor test (tab. 4) allowed rejecting zero hypothesis on the lack of differences in the effect of different photopolymerization procedures with regards to three materials: FSi, GDA and GDP. The largest differences have been observed in group of GDA results (F = 8.58). There have been no significant differences in comparisons of H material strength groups.

In direct comparisons of result groups in HSD Tukey's test (Tab. 5) there were no significant differences in strength of FSi material observed in three cases:

- between HAL 40 and HAL 60 groups,
- between HAL 60 and LED 40,
- between HAL 60 and LED 60.

Table 4. Variance analysis results (F – Snedecor test)

Analysis of Variance (p < 0,05)								
Material	***SS – Effect***	***df - Effect***	***MS – Effect***	***SS - Error***	***df - Error***	***MS – Error***	***F***	***p***
FSi	1968,09	3	656,03	9408,01	76	123,79	5,30	0,002266
GDA	1546,80	3	515,60	4567,62	76	60,10	8,58	0,000056
GDP	718,22	3	239,40	4880,74	76	64,22	3,73	0,014753
H	608,14	3	202,71	13321,91	76	175,29	1,16	0,332002

Table 5. Post-hoc HSD Tukey's test results (p < 0,05)

FSi				
	{1} - M=108,14	***{2} - M=119,50***	***{3} - M=108,25***	***{4} - M=107,76***
HAL 40s {1}		0,009864	0,99999	0,999588
HAL 60s {2}	0,009864		0,0108	0,007141
LED 40s {3}	0,99999	0,0108		0,999066
LED 60s {4}	0,999588	0,007141	0,999066	
GDA				
	{1} - M=77,765	***{2} - M=74,980***	***{3} - M=68,155***	***{4} - M=67,475***
HAL 40s {1}		0,66862	0,001193	0,000538
HAL 60s {2}	0,66862		0,033632	0,015878
LED 40s {3}	0,001193	0,033632		0,992558
LED 60s {4}	0,000538	0,015878	0,992558	
GDP				
	{1} - M=84,145	***{2} - M=82,030***	***{3} - M=79,660***	***{4} - M=76,075***
HAL 40s {1}		0,837899	0,295872	0,011169
HAL 60s {2}	0,837899		0,786106	0,096001
LED 40s {3}	0,295872	0,786106		0,494301
LED 60s {4}	0,011169	0,096001	0,494301	
H				
	{1} - M=107,22	***{2} - M=99,600***	***{3} - M=102,61***	***{4} - M=102,04***
HAL 40s {1}		0,271535	0,68929	0,60542
HAL 60s {2}	0,271535		0,88935	0,936703
LED 40s {3}	0,68929	0,88935		0,999175
LED 60s {4}	0,60542	0,936703	0,999175	

Additionally, significant differences have been demonstrated in strength of GDA material while comparing the following results groups: HAL 40 and LED 40, HAL 40 and LED 60, HAL 60 and LED 40, as well as HAL 60 and LED 60. In case of GDP material significant differences were shown by comparing groups HAL 40 and LED 60, as well as HAL 60 and LED 60. In case of H material there were no significant differences in strength (resulting from photopolymerization procedure).

In table 6 the results of approximation of experimental data by means of Weibull distribution are given: coefficient of determination R^2, characteristic strength (scale parameter) σ_0 and Weibull modulus (shape parameter) m. Average m parameter values and standard deviation are also shown in sets assigned to particular materials: FSi, GDA, GDP, H.

The highest Weibull modulus was obtained for result group GDP HAL 40 ($m = 27,28$). The highest average value of modulus was indicated for GDP composite ($m = 15,69$), however in this case also the largest standard deviation was observed (*st.dev. = 8,45*). It confirms a considerable diversification of the results between the photopolymerization subgroups. For FSi material the lowest deviation of the average m modulus value *(st.dev. = 2,38)* was noticed. The lowest average value of Weibull modulus was obtained for H material ($m = 9,00$; *st.dev. = 2,42*), also for this material a minimum modulus value of the whole studied population ($m = 5,88$) was demonstrated.

Table 6. Coefficient of determination R^2, characteristic strength σ_0 (MPa), Weibull modulus m of FSi, GDA, GDP, H composites in 3-point bending test

Material	**LCU**	**Time [s]**	σ_0	R^2	**Weibull mod. *m***	**Weibull mod. *m***	
						Mean	***St.Dev.***
FSi	HAL	40	112	0,97	14,25	12,31	2,38
		60	122	0,95	14,46		
	LED	40	112	0,93	10,65		
		60	112	0,93	9,88		
GDA	HAL	40	81	0,98	16,82	11,63	3,90
		60	81	0,97	10,24		
	LED	40	72	0,93	11,90		
		60	72	0,90	7,54		
GDP	HAL	40	86	0,96	27,28	15,69	8,45
		60	85	0,97	16,00		
	LED	40	84	0,97	11,88		
		60	79	0,94	7,60		
H	HAL	40	116	0,95	8,91	9,00	2,42
		60	106	0,86	5,88		
	LED	40	105	0,97	11,76		
		60	106	0,93	9,44		

A linear approximation after logarithmic transformation of the flexural strength test results, regression equations and coefficients of determination are shown in Figures 3 – 6.

4. Discussion

A varied effect of photopolymerization procedures depending on the type of composite on the flexural strength of the studied materials and their reliability has been shown. In a group of methacrylate based composites the average strength in the tested specimen (tab. 3) decreases in the statistically significant manner after switching from halogen (HAL) to LED lamp in case of GDA and GDP materials, whereas in case of H material changes are insignificant (tab. 4). It has been indicated that H is characterized by rather large scatter of results of each observation (variation coefficient 9.90% – 16.68%). A similar tendency of strength reduction has been also noticed while increasing exposure time from 40 to 60 seconds.

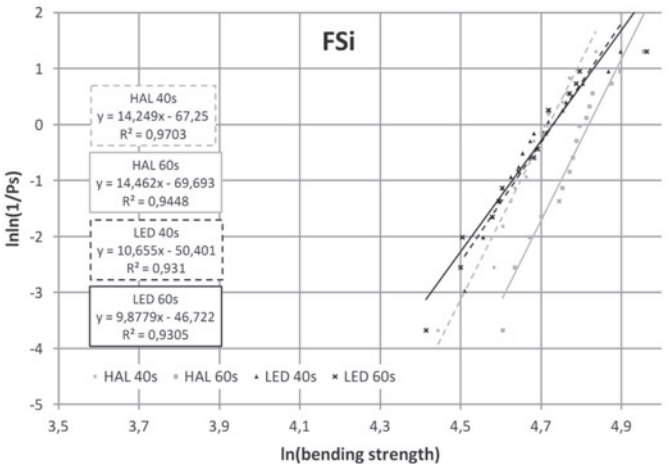


Fig. 3. Aproximation of probability distribution of flexural strength of FSi composite

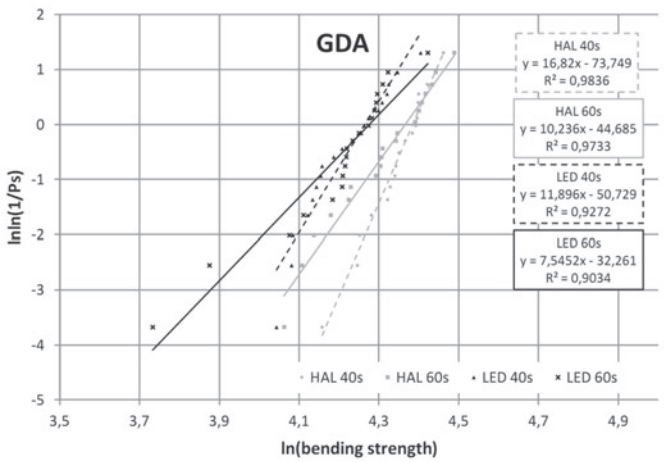


Fig. 4. Aproximation of probability distribution of flexural strength of GDA composite

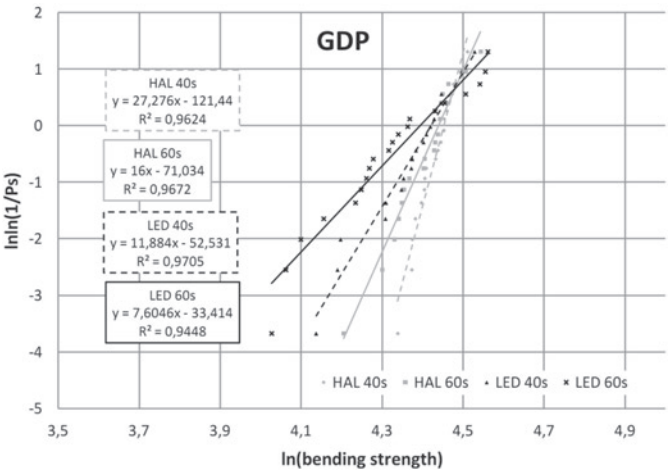


Fig. 5. Aproximation of probability distribution of flexural strength of GDP composite

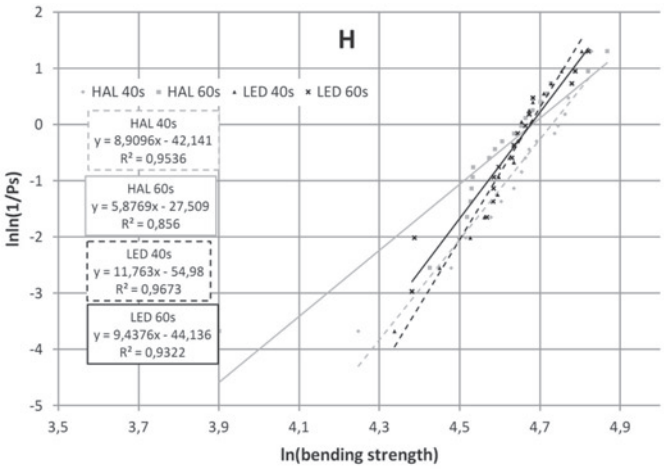


Fig. 6. Aproximation of probability distribution of flexural strength of H composite

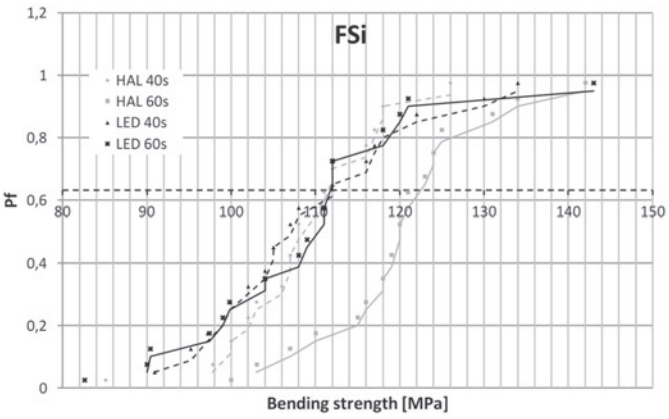


Fig. 7. Unreliability function of FSi composite with regards to flexural strength at different photopolimerization procedures

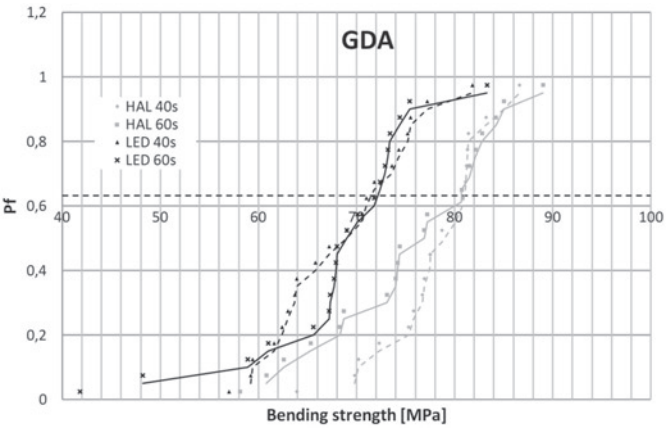


Fig. 8. Unreliability function of GDA composite with regards to flexural strength at different photopolimerization procedures

For the studied group of conventional polymers based on methacrylates, regularities described above are also valid in the aspect of reliability (tab. 6, fig. 8, fig. 9, fig. 10), although it is not so unequivocal. Switching from halogen to LED lamp has a clear and negative impact on the characteristic strength only in case of GDA material. For GDP material this effect is insignificant. In both cases the impact of prolonged exposure time has not been noticed. In case of H material, both changing the lamp as well as extending the exposure time, reduced the characteristic flexural strength. It is worth noticing that the Weibull modulus *m* for H material has a low value. In the literature the low Weibull modulus is interpreted as a result of large scatter of structural defects in the specimens, equal to its low reliability [20, 23].

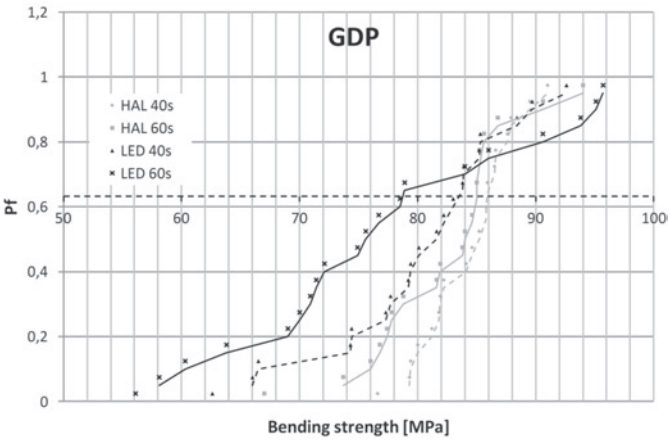


Fig. 9. Unreliability function of GDP composite with regards to flexural strength at different photopolimerization procedures

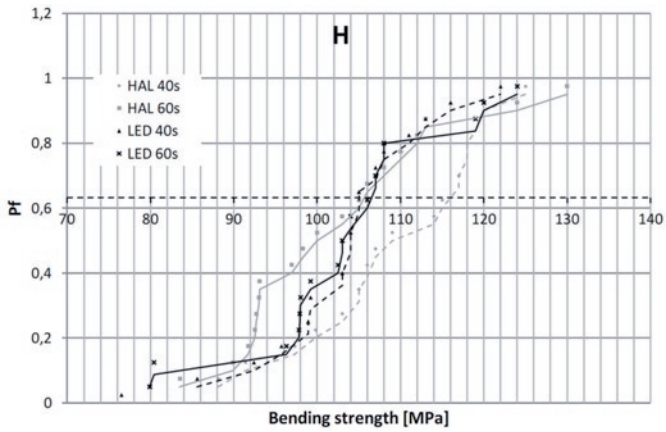


Fig. 10. Unreliability function of H composite with regards to flexural strength at different photopolimerization procedures

In comparison with conventional methacrylates-based composites, FSi material based on siloranes can be favorably distinguished. A high characteristic strength value (the highest of all studied materials) $\sigma_0 = 112$ to 122 MPa (tab. 6) and high Weibull modulus value $m = 12.31$ were observed, which gives a basis for a quite high rating of reliability of this material. Unreliability function curve (fig.7) indicates a rapid increase of failure probability due to flexural strength occurs above 100 MPa, while for conventional composites it occurs already at 60–90 MPa. It can result from difference in polymerization shrinkage, which as well know, is a cause of initiation of internal residual stresses in the material. The polymerization shrinkage of Filtek Siloran material is estimated at ≤ 1.0 % [12], while polymerization shrinkage of methacrylates-based composites is in a range of 2.0 – 3.0 % [29,33]. Additionally, it has been indicated that in case of FSi composite the extension of exposure time by halogen lamp from 40 to 60 seconds results in the increase of characteristic strength, while it has no impact on reliability (Weibull modulus). It has been demonstrated that switching to LED lamp has now impact on the strength of FSi composite.

The conducted studies have indicated that for some testing groups (GDA material – group LED 40 and LED 60, GDP material – group LED 60, H material – group HAL 60 and LED 60) there are low values of coefficient of determination (below 0.95). Some authors suggest that it can be caused by heterogeneity of the material properties in specimen as a result of imperfect photopolymerization procedure, consisting in overlapping exposure of the specimens surface by the light beam of the lamp [19, 28], thus producing double exposed areas in the specimens.

5. Conclusions

1. It has been demonstrated that in 3-point bending test conditions a change of the type of photopolymerization lamp from halogen (HAL) to LED brings the following results:
   - in case of FSi composite with low polymerization shrinkage it has no impact on the characteristic composite strength, while it has an impact on the increase of its reliability (specified by Weibull modulus);
   - in case of conventional composites based on methacrylate compounds (GDA, GDP, H) a reduction of characteristic strength occurs, while reliability of the material increases (increase of Weibull modulus).
2. It has been concluded that the extension of polymerization time from 40 to 60 seconds:
   - in case of Filtek Siloran (FSi) with use of halogen lamp effects in the increase of strength, however it does not change composite reliability. In case of LED lamp it has no impact on the strength and only insignificantly increases reliability;
   - in case of conventional composites based on methacrylates (GDA, GDP, H) with the use of both halogen and LED lamp has no impact on the strength, however it significantly improves material reliability.
3. It has been concluded that the applied method for analysis of the study's results, consisting in application of Weibull distribution relating to 3-point bending test, expends possibilities of preliminary assessment of operational usability of the new dental materials.

References

1. Asmussen E, Peutzfeldt A. Influence of UEDMA, Bis-GMA and TEGDMA on selected mechanical properties of experimental resin composites. Dental Materials 1998;14: 51–6.
2. Davies D.G.S. The statistical approach to engineering design in ceramics. Proceedings of the British Ceramic Society 1973;22: 429–52.
3. Della Bona A. Characterizing ceramics and the interfacialadhesion to resin: I – The relationship of microstructure composition, properties and fractography. J Appl Oral Sci 2005; 13: 1–9.
4. Della Bona A, Anusavice K J, DeHoff P H. Weibull analysis and flexural strength of hot-press core and veneered ceramic structures. Dent Mater 2003; 19: 662-669.
5. Dunn W J, Bush A C. A comparison of polymrization by light miting diode and halogen-based light curing units. J Am Dent Assoc 2002; 122: 335–341
6. Ferracane J L, Berge H X, Condon J R. In vitro aging of dental composites in water—effect of degree of conversion, filler volume, and filler/matrix coupling. Journal of Biomedical Materials Research 1998; 42: 465–72.
7. FiltekTM Silorane; www.3M.com.
8. Heath M R, Prinz J F. Oral processing of foods and the sensory evaluation of texture, in: A.J. Rosenthal (Ed.), Food Texture: Measurement and Perception, Gaithersburg, 1999.
9. ISO 4049 Dentistry – Polymer-based filling, restorative and luting materials 2000.

10. Kelsey W P, Latta M A, Shaddy R S, Stansilav C M. Physical properties of three packable resin-composite restorative materials. Operative Dentistry 2000; 25: 331–5.
11. Konopka T. Wprowadzenie do metodologii badań nieeksperymentalnych. Czas. Stomatol. 2009; (62)7: 597–604.
12. Lien W, Vandewalle K S. Physical properties of a new silorane-based restorative system. Dent. Mater. 2010; 26: 337–344.
13. Lodhi T.A. Surface hardness of different shades and types of resin composite cred with a high Power LED light curing unit. University of Western Cape; 2006.
14. Manhart J, Kunzelmann K H, Chen H Y, Hickel R. Mechanical properties of new composite restorative materials. Journal of Biomedical Materials Research 2000; 53: 353–361.
15. McCabe J F, Carrick T E. A statistical approache to the mechanical testing of dental materials. Dent. Mater. 1986; 2: 139–142.
16. Migdalski J. Inżynieria niezawodności. Poradnik. Wyd. ATR ZETOM, Warszawa 1992.
17. Moszner N, Salz U. New development of polymeric dental composites. Prog. Polym. Sci 2001; 26: 535–536.
18. Musanje L, Ferracane J L. Effects of resin formulation and nanofiller surface treatment on the properties of experimental hybryd resin composite. Biomaterials 2004; 25: 406–571.
19. Palin W M, Fleming G J P, Marquis P M. The reliability of standardized flexure strength testing procedures for a light-activated resin-based composite. Dental Materials 2005; 21: 911–919.
20. Palin W M, Fleming G J P, Burke F J T, Marquis P M, Randall R C. The reliability in flexural strength testing of a novel dental composite. J of Dent. 2003; 31: 549–557.
21. Peris A R., Mitsui F H O, Amaral C M, Ambrosano G M B , Pimenta L A F. The effect of composite type of microhardness when using quarto-tungsten-halogen (QTH) of LED lights. Oper Dent 2005; 30(5): 649–654.
22. Peutzfeld A. Resin composites in dentistry the monomer system. Eur J Oral Sci 1997; 105: 97–116.
23. Pick B, Meira J B C, Driemeier L, Braga R R. A critical view on biaxial and short-beam unaxial flexural strength tests applied to resin composites Rusing Weibull, fractographic and finite element analyses. Dent Mater 2010; 26: 83–90.
24. Pieniak D, Niewczas A M, Kordos P. Influence of thermal fatigue and ageing on the microhardness of polimer-ceramic composites for biomedical applications. Maintenance and Reliability – Ekploatacja i Niezawodnosc 2012; (14)2: 181–188.
25. Powers J M, Sakaguchi R L. Craig's Restorative Dental Materials, Twelfth Edition, 2006.
26. Ritter J E. Critique of test methods for lifetime predictions. Dent. Mater. 1995; 11: 147–151.
27. Rodriguez S A Jr, Ferracane J L, Della Bona A. Flexural strength and Weibull analysis of a microhybrid and a nanofill composite evaluated by 3- and 4- point bending tests. Dent Mater 2008; 24: 426–431.
28. Sakaguchi R L, Peters M C R B, Nelson S R, Douglas W H. Poort H W. Effects of polymerisation contr action in composite restorations. Journal of Dentistry 1992; 20: 178–82.
29. Santerre J P, Shaji Z, Leung B W. Relation of dental composite formulations to their degradation and release of hydrolyzed polymeric-resin-derived products. Crit Rev Oral Biol Med 2001; 12: 136-151.
30. Sinval A. Rodrigues Junior, Jack L. Ferracane, Alvaro Della Bona. Flexural strength and Weibull analysis of a microhybrid and a nanofill composite evaluated by 3- and 4-point bending tests. Dental Materials 2008; 24: 426–431.
31. Stanley P, Fessler H, Sivil A D. An engineer's approach to the prediction of failure probability in brittle components. Proceedings of the British Ceramic Society 1973; 22: 453–87.
32. Watts DC, Hindi A A. Intrinsic 'soft start' polymerisation shrinkage kinetics in an acrylate-based resin-composite. Dental Materials 1999; 15: 39–45.

Agata M. NIEWCZAS, DMD Ph.D.
Department of Conservative Dentistry
Medical University of Lublin,
Karmelicka 7 Str., 20-081 Lublin,
e-mail: agatan117@wp.pl

Daniel PIENIAK, Ph.D. (Eng.)
Department of Applied Mechanics,
Main School of Fire Service, Warsaw
52/54 J. Słowackiego Str., 01-629 Warsaw
e-mail: dpieniak@sgsp.edu.pl

Paweł OGRODNIK, Ph.D. (Eng.)
Department of Applied Mechanics,
Main School of Fire Service, Warsaw
52/54 J. Słowackiego Str., 01-629 Warsaw
e-mail: pogrodnik@sgsp.edu.pl

Article citation info:
ZHANG T, CHENG Z, LIU YJ, GUO B. Maintenance scheduling for multi-unit system: a stochastic Petri-net and genetic algorithm based approach. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (3): 256-264.

Tao ZHANG
Zhijun CHENG
Ya-jie LIU
Bo GUO

MAINTENANCE SCHEDULING FOR MULTI-UNIT SYSTEM: A STOCHASTIC PETRI-NET AND GENETIC ALGORITHM BASED APPROACH

USTALANIE HARMONOGRAMU OBSŁUGI DLA SYSTEMU WIELOELEMENTOWEGO: PODEJŚCIE OPARTE NA STOCHASTYCZNYCH SIECIACH PETRIEGO ORAZ ALGORYTMIE GENETYCZNYM

Frequent maintenance activities would cause low system availability and require large sums of money. For a multi-unit system, maintenance activities of some units can be combined together to reduce the total maintenance possession time and cost. Therefore, an optimized timetable of the maintenance activities is needed to be planned. Considering the uncertainties in both the deterioration and maintenance process of the units in a system, this paper advances a stochastic Petri-net based simulation optimization model for maintenance scheduling. The genetic algorithm is used to get the solution of the timetable of the maintenance activity schedule such that the overall cost is minimized in a planning horizon taking into account total maintenance possession time, unit condition, life cycle loss and solution feasibility. Some techniques used to reduce the computational effort required to perform the analysis are also described. A case study is given in the end.

Keywords: *Maintenance scheduling; multi-unit system; Petri net; genetic algorithm; deterioration; minimal cut set.*

Częste czynności obsługowe prowadzą do niskiej gotowości systemu oraz wymagają dużych nakładów pieniężnych. W systemie wieloelementowym całkowity czas i koszt obsługi można obniżać łącząc ze sobą czynności obsługowe niektórych elementów. Dlatego też konieczne jest planowanie zoptymalizowanego harmonogramu czynności obsługowych. W artykule zaproponowano model symulacyjny optymalizacji harmonogramu obsługi oparty na stochastycznych sieciach Petriego uwzględniający niepewność zarówno procesu deterioracji jak i procesu obsługi elementów systemu. Algorytm genetyczny wykorzystano do opracowania terminarza czynności obsługowych, który pozwalałby na minimalizację kosztów całkowitych w przyjętym horyzoncie planowania przy uwzględnieniu całkowitego czasu obsługi, stanu elementów, strat wynikających z cyklu życia oraz wykonalności rozwiązania. Ponadto opisano techniki zastosowane w celu zmniejszenia wysiłku obliczeniowego potrzebnego do wykonania analizy. W końcowej części pracy przedstawiono studium przypadku.

Słowa kluczowe*: Ustalanie harmonogramu obsługi; system wieloelementowy; sieć Petriego; algorytm genetyczny; minimalny przekrój niezdatności.*

1. Introduction

Most of the equipments are complex function-integration system, of which the deterioration and failures might incur high costs (e.g due to production losses and delays, service interrupt, unplanned intervention on the system) and safety hazards (e.g if the resistance of a deteriorated structure drops below the applied stress). So there has been a growing interest in the modelling and optimisation of maintenance of multi-unit system which means the units of system depend on each other (i.e., economic/stochastic dependence) [1]. Economic dependence [2-3] implies either cost can be saved when several units are jointly maintained instead of separately, whereas stochastic dependence means that each component's transition probability depends on the other components'. Then optimal maintenance policies for such systems cannot reduce to those for systems with a single unit. A decision must be made to improve the whole system, rather than any subsystem.

The maintenance scheduling problem of multi-unit system has been researched and surveyed by several people. Cho and Palar [1] gives an overview of the multi-unit maintenance literature up to 1991, including machine-interference/repair models, group-replacement models of various types, spare-parts models, and inspection models. Dekker et al. [2] exclusively deals with multi-component maintenance models based on economic dependence. Later [4] presents another paper that surveys this field by different category ways. Furthermore, Wang [3] reviews the maintenance policies of deteriorating systems, of which one section is devoted to opportunistic maintenance policies for multi-unit systems with economic dependence. From these reviews we can find the maintenance model of multi-unit system is too complex to solve, especially considering two dependencies or more, so only the economical dependence is discussed in most references [2,6]. Generally there are two important policies to organize the maintenance of multi-unit system with economical dependence. One is group maintenance [5] under which the system is either entirely replaced with new components or is allowed to remain in operation. The other is opportunistic maintenance [6] under which preventive maintenance is carried out at opportunities, either by choice or based on the physical condition of the system. Most of the existing group/opportunistic maintenance models allow grouping of the maintenance tasks, but few of them are proposed in the context of the condition-

based maintenance [7], especially when the number of units increases. The mathematical formulations of these maintenance models become too complicated to get the analytical solution. Therefore, many papers present the simulation method to solve the problem. Ouali [8] proposes a simulation model for opportunistic maintenance strategies. Preventive maintenance activities are combined with corrective one by using Promodel software program. E.Zio [9,10] discussed the simulation method to get the optimal opportunistic policies under several kinds of condition. However, most of literatures [11,12] only focus on the optimisation technique and seldom discuss the influence and allocation of resources.

In the case of a continuous deteriorating process and limited resources in maintenance, we propose, in this paper, three novel developments when compared with the previous work on multi-unit maintenance modeling and optimisation. First, the deterioration model is presented which allow us to investigate the uncertainty of the deterioration and maintenance process. The assumption will not limit the type of probability distribution any more. Second, multi-type resources and the behaviour of competing and sharing can be considered in a stochastic Petri-net (SPN) based model [13]. Finally, a GA-based approach is advanced to find a satisfactory solution. Such approach has been successfully applied to many engineering optimization problems including maintenance scheduling.

The rest of the paper is organized as follows. Section 2 gives the description of this problem, the framework of our approach, the objective function and how to compute the total maintenance possession time. Section 3 presents the SPN-based model for calculating the objective. Section 4 describes the steps of genetic algorithm which is used for searching an optimal solution. Section 5 presents an example of a multi-unit system with 10 units. The final section makes a conclusion.

2. Model formulation

2.1. Problem description

Most units in a system will not fail suddenly but deteriorate from a good condition to an unacceptable condition. This study divided the deterioration process into three phases; good condition, trigger condition and unacceptable condition, as shown in Figure 1. The trigger condition is the deterioration point that can be detected by existing methods and devices. But each unit would be functional in the trigger condition until it deteriorates to the unacceptable condition. Normally, the interval between trigger and unacceptable condition, called the time window, is uncertain. We may assume that it follows a kind of probability distribution based on condition history data analysis.

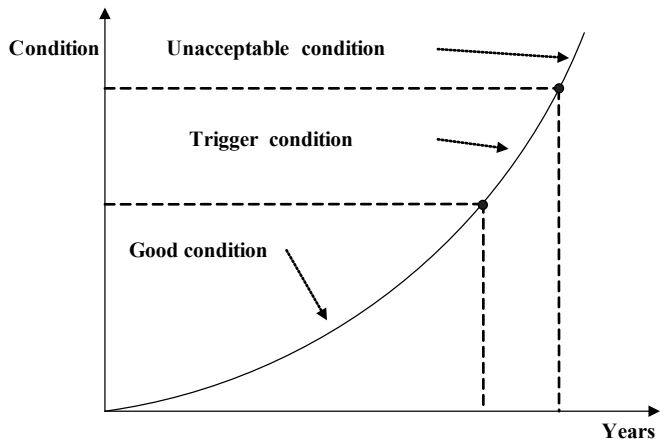


Fig. 1. The deterioration process

It is not possible to wait until the unit deteriorates to an unacceptable condition because it may be very dangerous when it is working in this condition and the cost for interruption of allowing a failure will be huge. Hence, it would be better to repair or renew the units before they deteriorate to an unacceptable condition resulting in system down. So the time window is the best time to start maintenance activity. On the other hand, a long term timetable for maintenance or renewal activities should be planned in order to leave more time for users to arrange the operation schedule.

The problem in this study is that there is a multi-unit system

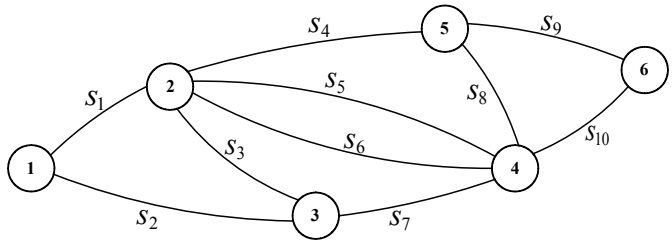


Fig. 2. The transmission network system with 10 units

$S=(S_1,S_2,\ldots,S_N)$. N is the number of units and S_p represents the pth unit, $p=1,2,...,N$. Let $F=f(S_1,S_2,\ldots,S_N)$ be the system failure logic function. For example, there is a transmission network system with 10 units, where the source and sink node respectively is 1 and 6. Hence, we can get

$$F = S_1S_2 + S_1S_3S_7 + S_2S_3S_4S_5S_6 + S_4S_5S_6S_7 + S_1S_3S_5S_6S_8S_{10} + \\ + S_2S_3S_5S_6S_8S_9 + S_5S_6S_7S_8S_9 + S_4S_8S_{10} + S_9S_{10} \quad (1)$$

The conditions of different units are assumed to be statistically independent. All units are considered to be repaired or renewed in a planning horizon H, the time windows of them are not determinate and is described by a probability distribution, the maintenance resources are limited, the aim is to give an optimal schedule for maintenance activities of N units in a finite horizon taking into account the condition of units, the maintenance possession time, the life loss and the plan feasibility.

2.2. Framework of our approach

There are two key problems that should be solved. One is how to quantify and calculate the benefit of a solution. Another is how to find the best solution.

For the first problem, in this study, the failure time and maintenance time of different units in a system are characterized by different probability distributions. There are also resources shared in the maintenance process. A total cost is defined to quantify the benefit of a solution. And because the SPN is suitable for describing the behaviour of a dynamic system, it is used in this approach for analyzing the system dynamic behaviour and computing the objective of the given solution by simulation.

For the optimization problem, it is known that the solution varies non-linearly with the continuous variables, and that the size of the problem dramatically increases with the number of units considered. Such a scheduling problem is proved to be an NP-hard problem. The standard methods of non-linear programming are not suited particularly when the number of units considered is large. The GA-based approach has been successfully applied to many engineering optimization problems including maintenance scheduling. Hence, GA is also used to solve this problem.

Figure 3 shows the framework of the approach used. The SPN and its simulation are described in Section 3, steps of the GA are described in detail in Section 4.

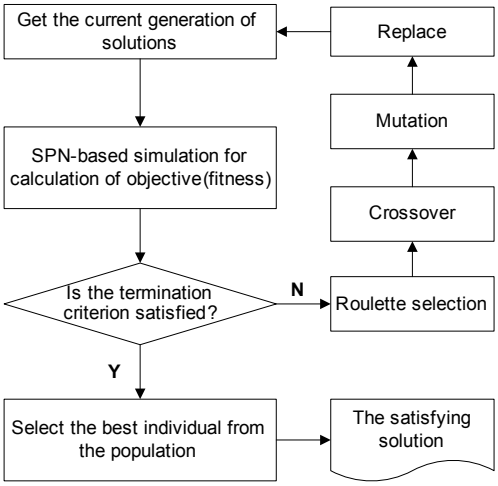


Fig. 3. The framework of SPN and GA-based approach

2.3. The objective function

In this study, the solution is described by the composition of planned maintenance start time of *N* units in a system given. Let T_p be the planned start time of the maintenance activity for the pth unit, $0 < T_p < H$, $p = 1,2,...,N$. So a schedule solution can be described by the vector of $(T_1, ..., T_p, ..., T_N)$. An example is presented in Figure 4, there are five units considered in the planning horizon. Here $(T_1, T_2, T_3, T_4, T_5)$ is a solution vector.

To quantify the benefit of a solution, the total cost was defined taking into account four factors in this study.

First, to reduce cost caused by system service interruption, the total maintenance possession time should be reduced. The total maintenance possession time T^{POSS} is a function of all actual maintenance time. If all units are in series, the total maintenance possession time would be sum of all actual maintenance time. However, all units are connected with a reliability logic relationship. The total maintenance possession time could not be given by simple sum. How to calculate possession time is given in section 2.4. Let c^{poss} be the possession cost per time unit. The total possession cost C^{POSS} is given by

$$C^{POSS} = c^{poss} T^{POSS} \tag{2}$$

Second, we should try to cut down the impact of unplanned maintenance activities. In this example, the 5th unit deteriorated to the unacceptable condition before the planned maintenance started. The maintenance activity of the 5th unit should start as soon as possible. The more time the unit is in the unacceptable condition, more dangerous the service is. So the actual start time will be before the planned one. Let T_p^m and T_p^u be the actual maintenance start time and of an unacceptable failure time of the pth unit respectively, $p = 1,2,...,N$.

Let c_p^{unac} be the cost per time unit of the pth unit in an unacceptable condition. Hence, the total cost due to any unit in an unacceptable condition C^{UNAC} is given by sum of that of all units as

$$C^{UNAC} = \begin{cases} \sum_{p=1}^{N} c_p^{unac} \left(T_p^m - T_p^u\right) & T_p^m > T_p^u \\ 0, & other \end{cases} \tag{3}$$

Third, if the time of maintenance for a unit is before the trigger time as a result of it being combined with the maintenance of other units, its life cycle loss may be increased. So it would be better to start the maintenance activity of units in their time windows. In this example, the maintenance activities of the 2nd unit started before the trig-

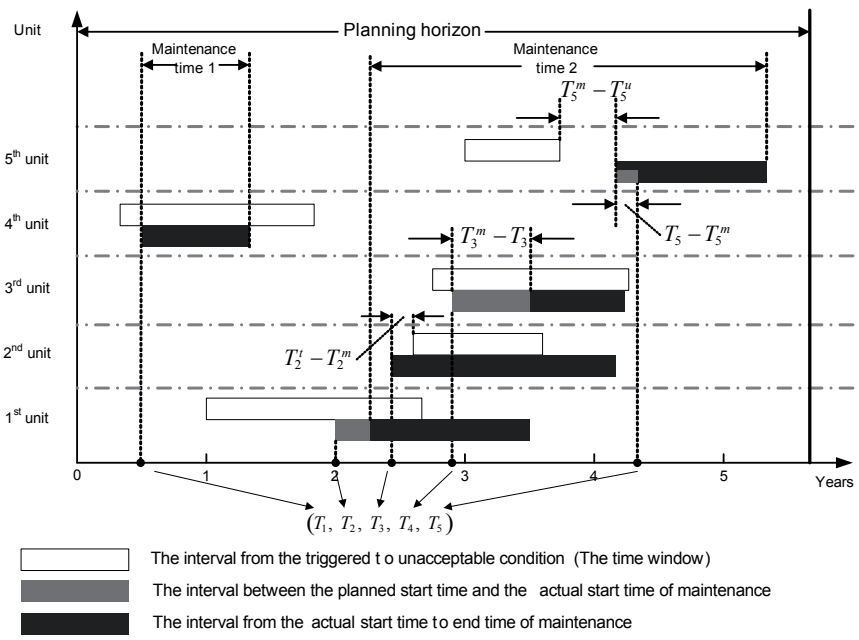


Fig. 4. An example of the maintenance activities of five units

ger time. Let T_p^t be the trigger time of the pth unit and c_p^{window} be the penalty cost per time unit for the advanced time before the trigger time of the pth unit. The penalty cost C^{WIND} caused by the opportunistic maintenance activities starting before the trigger time is given by:

$$C^{WIND} = \begin{cases} \sum_{p=1}^{N} c_p^{window}\left(T_p^t - T_p^m\right) & T_p^t > T_p^m \\ 0, & other \end{cases} \tag{4}$$

Fourth, because of the limit of resources and the uncertainties of the deterioration and maintenance processes, the maintenance activity may not be able to start at the planned time. In the example of Figure 4, the maintenance activities of the 1st, 3rd and 5th units didn't start at the planned time. However, we should try to decrease the waiting time for maintenance resources and guarantee the feasibility of solution. Let c_p^{plan} be the penalty cost per time unit for the activity of the pth unit couldn't start at the planned time. The penalty cost C^{PLAN} caused by the maintenance activities can't start at planned time is given by:

$$C^{PLAN} = \sum_{p=1}^{N} c_p^{plan}\left|\left(T_p - T_p^m\right)\right| \tag{5}$$

where $\left|\left(T_p - T_p^m\right)\right|$ is absolute value of $\left(T_p - T_p^m\right)$, it stands for interval time between the actual maintenance start time and the planned one.

Hence, the total cost is defined as:

$$C^{TOTAL} = C^{POSS} + C^{UNAC} + C^{WIND} + C^{PLAN} \tag{6}$$

The objective of maintenance scheduling, as described above, is to minimize the total cost by searching the best solution.

2.4. Calculation of the total maintenance possession time

Because all units are connected with each other in a system failure logic function described by *F*. The total maintenance possession time cannot be given by simple sum. In this study, the total maintenance possession time is defined as the total time that the system should be interrupted because of maintenance activities. In the system reliability problem, a cut set is a unit set such that if all units in the set are not in working condition, the system cannot work either. And a minimal cut set (MC) is an unit set such that if any unit is removed from the set, then the remaining set is no longer a cut set. For example, $\{S_1, S_3, S_7\}$ is an MC of the system shown in Figure 2. It is easy to know that if all units in any MC are scheduled to do maintenance activities in a same period, the system will be down. How to get all MCs of a system with failure logic function *F* is a classical problem in reliability analysis and there are many methods to solve it [14-15]. Hence, this study assume that all MCs are pre-computed, which are employed to compute the total maintenance possession time.

Let $\left(T_p^{Begin}, T_p^{End}\right)$ be the maintenance time of the pth unit, where T_p^{Begin} is the unacceptable condition begin time or the maintenance start time and T_p^{End} is the maintenance end time. Let $MCs = \{MC_1, MC_2, ..., MC_K\}$ be all MCs of the system. K is the number of MC and MC_i is the ith MC. The total maintenance possession time T^{POSS} can be calculated by:

$$T^{POSS} = \int_0^H f(t)dt \tag{7}$$

where

$$f(t) = \begin{cases} 1, & \exists MC_i \in MCs, MC_i \subseteq \left\{S_p \middle| T_p^{Begin} \le t \le T_p^{End}, p = 1,2,...,N\right\} \\ 0, & other \end{cases} \tag{8}$$

For the example shown in Figure 2, all MCs are $\{S_1, S_2\}$, $\{S_1, S_3, S_7\}$, $\{S_2, S_3, S_4, S_5, S_6\}$, $\{S_4, S_5, S_6, S_7\}$, $\{S_1, S_3, S_5, S_6, S_8, S_{10}\}$, $\{S_2, S_3, S_5, S_6, S_8, S_9\}$, $\{S_5, S_6, S_7, S_8, S_9\}$, $\{S_4, S_8, S_{10}\}$ and $\{S_9, S_{10}\}$. If the maintenance time of 10 units are respectively $(10,21)$, $(30,35)$, $(30,41)$, $(153,180)$, $(120,128)$, $(32,54)$, $(180,195)$, $(172,192)$, $(190,220)$, $(175,203)$. Here, the unit of time is a day. The fourth, eighth, and tenth units are not working for the unacceptable condition or maintenance in the 174th day. Hence, $f(174)=1$ because of existing an MC= $\{S_4, S_8, S_{10}\} \subseteq \{S_4, S_8, S_{10}\}$ and it stands for the whole system is not working. But in the 35th day, the second, third, and sixth units are not working and there does not exist an MC which is a subset of $\{S_2, S_3, S_6\}$, the system would be working, so $f(35)=0$. By formula (7), the total maintenance possession time is 20 in this example.

3. SPN-based simulation for computing the objective

Due to the uncertainties and the fact that the maintenance resources are shared in the maintenance process of the system, this problem is complex. The Petri-net method is suitable for describing the dynamic behaviour of the system. First created in 1962 and reported in the thesis of Petri[16], Petri-net are an adaptable and versatile, yet simple, graphical modelling tool used for dynamic system representation. In this study, a stochastic Petri-net is used for analyzing the system dynamic behaviour and computing the objective of a given solution by simulation.

3.1. The deterioration and maintenance process modelling

From section 2.1, it is known that the unit has three states – good condition, trigger condition and unacceptable condition. A SPN representing for the deterioration process including the maintenance process is shown in Figure 5. Three places P_0, P_1 and P_2 stand for three system states respectively, good condition, trigger condition and unacceptable condition. The location of the token indicates the state in which the system resides. The two transitions T_0 and T_1 respectively stand for the deterioration processes from good condition to trigger condition and from trigger condition to unacceptable condition. The transition duration time may follow any probability distribution, such as Weibull, exponential, gamma, normal, lognormal, beta or triangular, as given by the deterioration process of the unit. The place P_3 stands for the maintenance resources, such as maintenance teams,

workers, machines and tools. The initial quantity of tokens in place P_3 is decided by the number of available maintenance resources. The transaction T_2 stands for the maintenance process when the unit is in the trigger condition. The transaction T_3 stands for the maintenance process when the unit is in the unacceptable condition. If there are enough maintenance resources (enough tokens in Place P_3) when the unit is in the trigger condition, the transition T_5 will be triggered. But if there is not enough maintenance resource before the unit deteriorates to the unacceptable condition, the transition T_4 will be triggered after enough maintenance resources are ready.

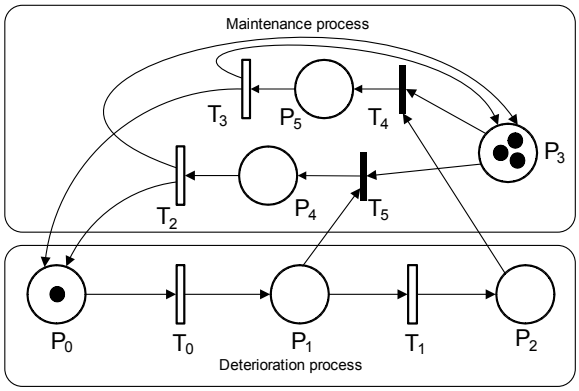


Fig. 5. The SPN of deterioration and maintenance process

3.2. SPN-based model

To calculate the objective function value under a given solution, some parameters need to be given. So we modify the SPN as shown in Figure 6 according to the basic model in Section 3.1. There are four parts in the Petri-net of one unit. The first part is the subnet of the deterioration process as described in Section 3.1. Initially, a token is in the place P_0. The second part is the subnet for maintenance scheduling. Initially, a token is in the place P_3. The transition time of T_4 is the planned start time of this unit and it is determined when the solution is given. When the token enters into the place P_4, it identifies that maintenance is required for this unit. The third part is the maintenance process. It is a little different from Figure 5. The priority of T_6 is higher than the priority of T_5 because the maintenance activity may be different after the unit gets to the unacceptable condition. If there is

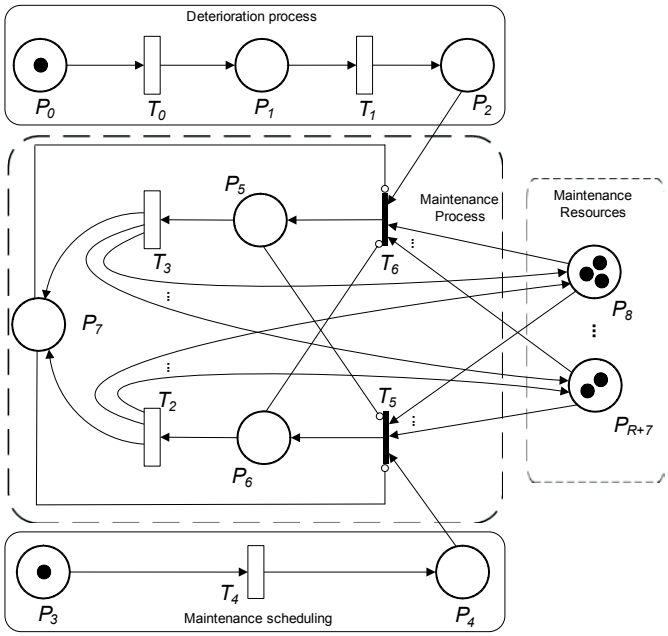


Fig. 6. The basic SPN of one unit for scheduling

a token in the place P_5 or P_7, transition T_5 can't be triggered because an inhibitor arc is activated. If there is a token in the place P_6 or P_7, transition T_6 can't be triggered too. It means only one transition of T_5, T_6 will be triggered. If both transitions T_5 and T_6 are satisfied, transition T_6 will be triggered because of its higher priority. In the fourth part is the subnet of maintenance resources, the places P_8, P_9,..,P_{R+7} stand for free maintenance resources of the 1st, 2nd,…,Rth type. R is the number of maintenance resource types. The number of tokens in these places stands for the number of available resources of the corresponding type.

Because of the assumption that any unit will not fail again after maintenance during the planning horizon, the token in P_7 will not leave again. If there are enough tokens in the places P_8,..,P_{R+7} when place P_2 or P_4 get a token, there are two scenarios of system behaviour. When the token arrival time of place P_4 (planned start time of the unit) is before that of place P_2 (the arrival time of unacceptable condition). Transition T_5, T_2 will be triggered and the scheduled maintenance activity will be done (scenario 1). Otherwise, the token arrival time of place P_4 is behind that of place P_2. Transition T_6, T_3 will be triggered and the maintenance activity caused by the unit in an unacceptable condition should be done (scenario 2). If there are not enough tokens in the resource places when place P_2 or P_4 get a token, there are also another two scenarios. When place P_2 gets a token before the resource places get enough tokens, transition T_6, T_3 will be triggered in any case when the resource places get enough tokens (scenario 3). When the resource places get enough tokens and place P_4 have a token but place P_2 have no token. Transition T_5, T_2 will be triggered when the resource places get enough tokens (scenario 4). These four scenarios suitably stand for four actual cases.

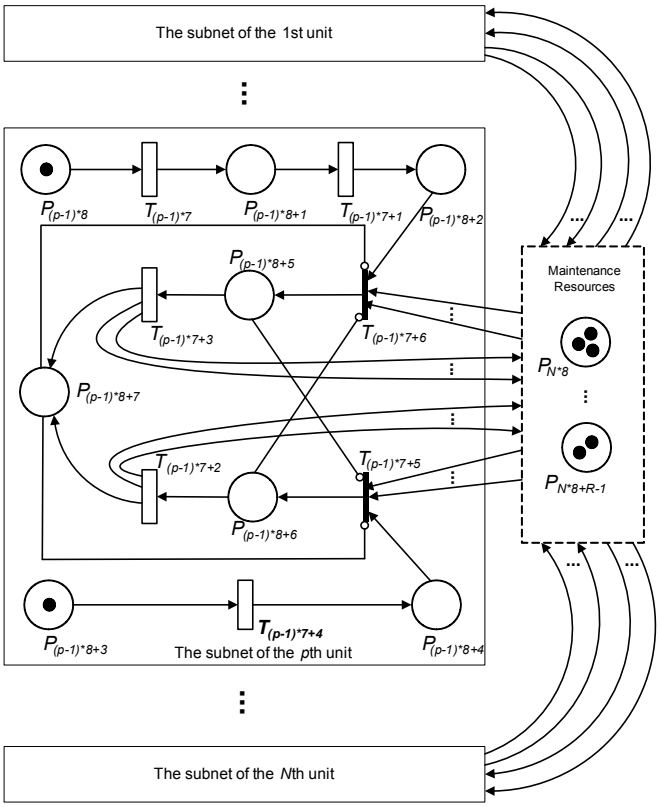


Fig. 7 The SPN of a system with N units

Based on the basic SPN of one unit for scheduling, we can easily build the SPN of the system by combining all SPNs of the considered units, as shown in Figure 7. The maintenance resources in places P_{N*8}, P_{N*8+1}, …, $P_{N*8+R-1}$ are shared by all units. The combination of the transition times of (T_4, …, $T_{(i-1)*7+4}$,…$T_{(N-1)*7+4}$) is the solution of the

maintenance schedule. The initial numbers of tokens in P_{N*8}, P_{N*8+1}, …, $P_{N*8+R-1}$ are the numbers of maintenance resources of the corresponding type respectively.

3.3. Calculation of the objective

Let S_N be the number of simulations. All values with subscript i stand for the values calculated in the ith simulation. For example, C_i^{TOTAL} stands for the total cost calculated in the ith simulation. So the average total cost is given by:

$$C^{TOTAL} = \frac{1}{S_N}\sum_{i=1}^{S_N} C_i^{TOTAL} = \frac{1}{S_N}\sum_{i=1}^{S_N}\left(C_i^{POSS} + C_i^{UNAC} + C_i^{WIND} + C_i^{PLAN}\right) \tag{9}$$

Let $T^{Token\_In}(i)$ be the token first arrival time of the ith place. This is obtained during the simulation. For scenarios 1 and 4 in section 3.2, the maintenance possession begin time of the pth unit T_p^{Begin} is the token arrival time of the $((p-1)\times 8+6)$th place. For scenario 2 and 3, it is the token arrival time of the $((p-1)\times 8+2)$th place. So,

$$T_p^{Begin} = \begin{cases} T^{Token\_In}((p-1)\times 8+6), & T^{Token\_In}((p-1)\times 8+6) > 0 \\ T^{Token\_In}((p-1)\times 8+2), & T^{Token\_In}((p-1)\times 8+2) > 0 \ and \ T^{Token\_In}((p-1)\times 8+6) = 0 \\ 0, & other \end{cases} \tag{10}$$

The maintenance possession end time of the pth unit T_p^{End} is the token arrival time of the $(p-1)\times 8+7$ th place. If $T_p^{Begin} \neq 0$ and no token arrives in the $(p-1)\times 8+7$ th place, it means that the maintenance activity can't be finished in the planning horizon. T_p^{End} is considered to be H. Hence,

$$T_p^{End} = \begin{cases} T^{Token\_In}((p-1)\times 8+7), & T^{Token\_In}((p-1)\times 8+7) > 0 \\ H, & T^{Token\_In}((p-1)\times 8+7) \leq 0 \ and \ T_p^{Begin} > 0 \end{cases} \tag{11}$$

By formula (6), the total possession time T_i^{POSS} and C_i^{POSS} can be calculated.

As shown in formulae (3), (4) and (5), to calculate C^{UNAC}, C^{WIND} and C^{PLAN}, we only need to get T_p^m, T_p^u and T_p^t. The actual maintenance start time of the pth unit T_p^m is the token arrival time of the $((p-1)\times 8+6)$th or $((p-1)\times 8+5)$th place. If no token arrives in these two places, it means that the maintenance activity can't be started in the planning horizon. We can set it to be H. So,

$$T_p^m = \begin{cases} T^{Token\_In}((p-1)\times 8+6), & T^{Token\_In}((p-1)\times 8+6) > 0 \ and \ T^{Token\_In}((p-1)\times 8+5) = 0 \\ T^{Token\_In}((p-1)\times 8+5), & T^{Token\_In}((p-1)\times 8+5) > 0 \ and \ T^{Token\_In}((p-1)\times 8+6) = 0 \\ H, & other \end{cases} \tag{12}$$

The unacceptable condition begin time of the pth unit T_p^u is when a token arrives in the $((p-1)\times 8+2)$th place. So,

$$T_p^u = \begin{cases} T^{Token\_In}((p-1)\times 8+2), & T^{Token\_In}((p-1)\times 8+2) \geq 0 \\ H, & Other \end{cases} \tag{13}$$

The trigger condition begin time of the pth unit T_p^t is when a token arrives in the $((p-1)\times 8+1)$th place. So,

$$T_p^t = \begin{cases} T^{Token\_In}((p-1)\times 8+1), & T^{Token\_In}((p-1)\times 8+1) \geq 0 \\ H, & Other \end{cases} \tag{14}$$

Hence, the objective could be calculated by the simulation of the SPN shown in Figure 7.

4. GA-based optimization approach

4.1 Genesis of the population

Planning horizon H is the maximum permitted value of each maintenance planned start time. To be dealt with easily, we use a vector of floating point numbers $(q_1, ..., q_p, ..., q_N)$ as an individual of the population also called a chromosome in GA terms, $0 < q_p < 1$, $p = 1, 2, ..., N$. The chromosome $(q_1, ..., q_p, ..., q_N)$ stands for the solution as below.

$$(T_1, ..., T_p, ..., T_N) = (q_1 H, ..., q_p H, ..., q_N H) \quad (15)$$

The initial population is created randomly. The population size should be large enough to search for an optimal solution. However, the larger the population size, the greater the computing time required.

4.2 Calculation of fitness

Using the SPN described in section 3, the objective of each individual, known as the fitness in the GA, was calculated by simulation. The bigger the number of simulation times, the more accurate the fitness becomes but the more searching time would be consumed. Hence, to improve the efficiency, a buffer and multi-thread programming technology was used. To avoid the repeated simulation of the same solution, the fitness value is saved in the buffer after each simulation. The fitness value would be used directly if the same chromosome appears again in a later generation. On the other hand, the generation would be divided into several groups before calculating their fitness. The simulation of different groups will be executed in different parallel threads. This method is called a multi-thread programming technology. The aim is to improve the execution efficiency.

4.3 Roulette selection

The roulette selection process is to choose the chromosomes to act as parents to perform crossover on. The crossover process is described in Section 4.4. After that, the next generation will be created. In this study, the probability of a chromosome being chosen is inversely proportional to its fitness describing the total cost. So the less the fitness, the greater chance the chromosome will be selected. Hence, the next generation is likely to be better than the previous one.

4.4. Crossover

The parent chromosome pair may crossover with a probability, known as the crossover rate P_c. This study used a one-point crossover process. It consists of selecting the crossover point p randomly, dividing each parent chromosome into two parts, swapping the corresponding part with the other chromosome in the pair to produce two new child chromosomes and replacing the parent chromosomes with child chromosomes. This process is shown as Figure 8.

Parent chromosome 1

$$(q_1^1, ..., q_{p-1}^1, q_p^1, q_{p+1}^1, ..., q_N^1)$$

Parent chromosome 2

$$(q_1^2, ..., q_{p-1}^2, q_p^2, q_{p+1}^2, ..., q_N^2)$$

$$\xrightarrow[p]{\text{Crossover}}$$

Child chromosome 1

$$(q_1^1, ..., q_{p-1}^1, q_p^2, q_{p+1}^2, ..., q_N^2)$$

Child chromosome 2

$$(q_1^2, ..., q_{p-1}^2, q_p^1, q_{p+1}^1, ..., q_N^1)$$

Fig. 8. Crossover

4.5. Mutation and replacement

In order to avoid the solution converging to a local best result, a mutation process is necessary. In this study, every value in the chromosomes may mutate with a probability, known as the mutation rate P_m. After mutation, the old population will be replaced by the new one.

5. Case study

This case study bases on the example shown in Figure 2. There is a system with 10 units that should be considered for renewal in the coming five years. Both the deterioration time from now to the trigger condition and the deterioration time from the trigger to unacceptable conditions follow the Weibull distribution. The renewal time follows the triangular distribution. Table 1 gives all distribution parameters of these units. The unit of time is a day.

In the SPN of this example, because we take into account the coming 5 years, the planning horizon was set to 1800 days. And for each solution, the simulation time was 1000. The parameters c^{poss}, c_p^{unac}, c_p^{window} and c_p^{plan} of all units were 400, 600, 300 and 1000 respectively.

In the GA-based searching for the optimal solution, it produced 30 generations and the population size of each generation had 200 individuals. The crossover rate P_c was 0.9 and the mutation rate P_m was 0.05. The number of teams is 4. Figure 9(a) and 9(b) respectively show the total cost curve of generation number in two cases; a transmission network system shown in Figure 2 and a series system. We can find that all the minimal, average and maximal total cost becomes lesser during the evolution process of the GA. Figure 10 and 11 respectively show the Gantt graphs of the optimal solution produced by the GA in two cases.

Table 1. Distribution parameters of the units

Seg. No.	Distribution of deterioration time from now to trigger condition	Distribution of deterioration time from trigger to unacceptable condition	Distribution of renewal time
1	Weibull, Scale=160, Shape=2	Weibull, Scale=200, Shape=2	Triangular. If renewal before the trigger condition, min=80, mean=100 and max=120 If renewal after the unacceptable condition, min=90, mean=120 and max=140
2	Weibull, Scale=210, Shape=2		
3	Weibull, Scale=220, Shape=2		
4	Weibull, Scale=500, Shape=2		
5	Weibull, Scale=650, Shape=2		
6	Weibull, Scale=700, Shape=2		
7	Weibull, Scale=600, Shape=2		
8	Weibull, Scale=950, Shape=2		
9	Weibull, Scale=1200, Shape=2		
10	Weibull, Scale=1230, Shape=2		

In these solutions, because there are four maintenance teams, the maintenance activities of some units whose deteriorated conditions are similar would be combined taking into account the opportunity to cut down the total maintenance possession time. When all units are not in series, the total maintenance posses-

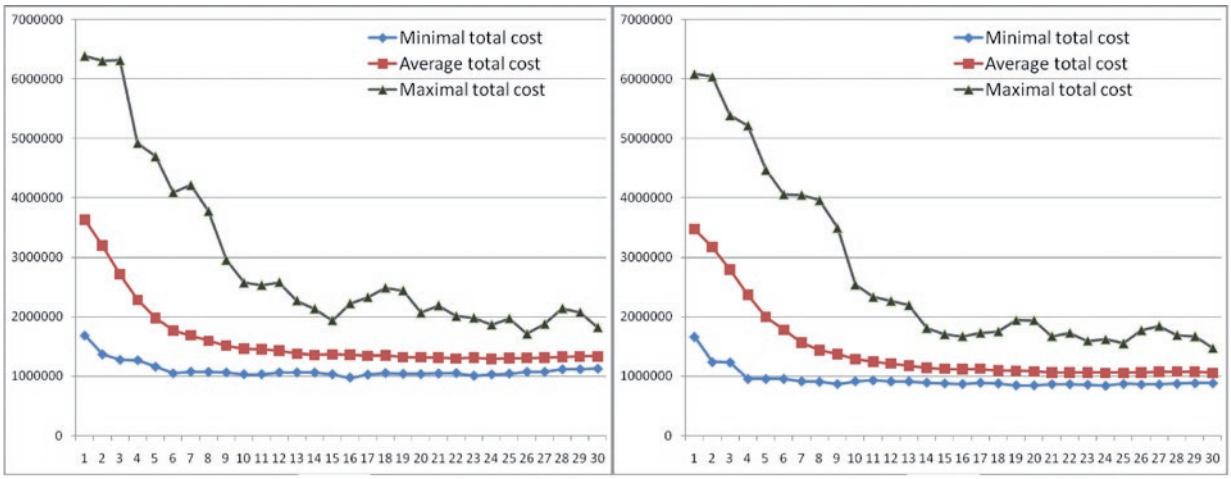


Fig. 9. The total cost curve of the generation number

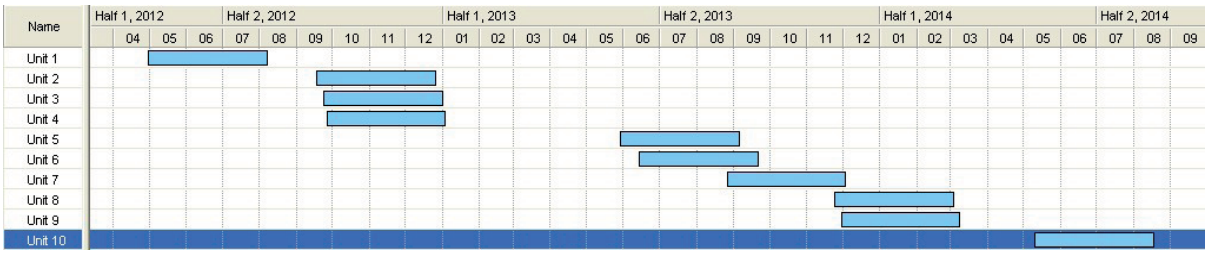


Fig. 10. The Gantt chart of the optimal solution of the system shown in Fig. 2 when the number of teams is 4

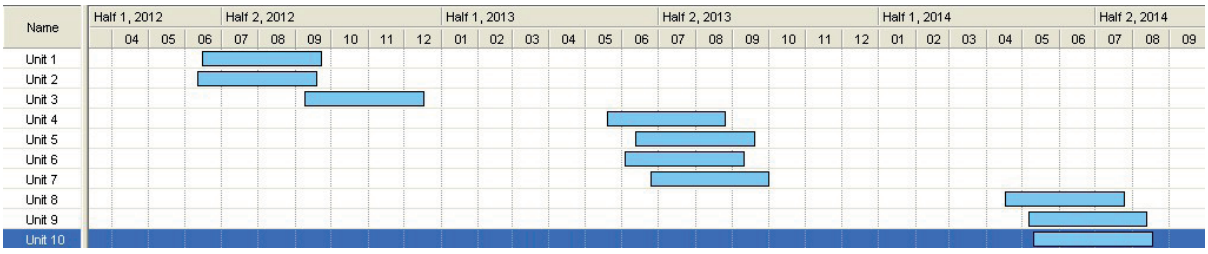


Fig. 11. The Gantt chart of the optimal solution of a series system when the number of teams is 4

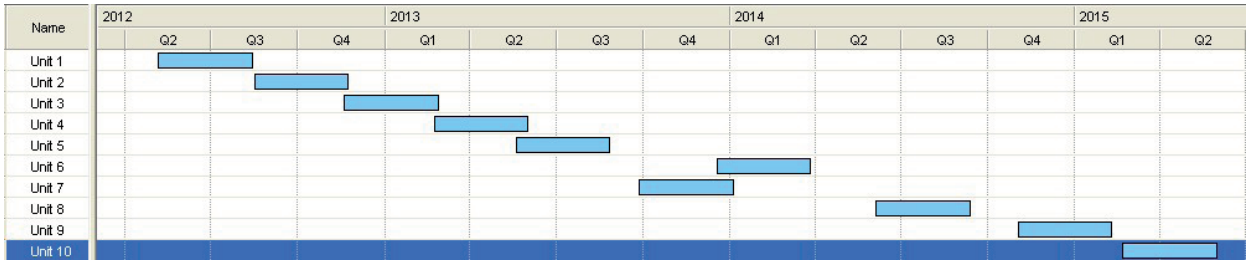


Fig. 12. The Gantt chart of the optimal solution when the number of teams is 1

sion time is not the simple sum of that of all units. It has to be avoided that the units set whose maintenance activities are in a same period belongs to a cut set. Hence, although the conditions of unit 1 and 2 are very close, it is better not to combine the maintenance activities of them. Analogously, it is better not to combine the maintenance activities of the unit 9 and 10. It is different from the case when all units are in series.

Figure 12 shows the Gantt graphs of the optimal solution as produced by the GA when there is only one available maintenance team. All maintenance activities would be scheduled in series because of the limit of the maintenance teams. The worse the condition of the component is, more chance to be scheduled in advance

6. Conclusions

This paper describes an approach for maintenance scheduling of a multi-unit system in a finite planning horizon. In this approach, SPN was used to describe the stochastic and dynamic behaviour of the deterioration and maintenance process of the units. It evaluates the total cost including maintenance possession cost, life cycle loss and penalty cost as the objective which can be calculated by the simulation of the SPN. Because this kind problem was proved to be an NP-hard problem, so the GA was investigated to search for the optimal solution. To improve the efficiency of the GA, a buffer and multi-thread programming technology were used to avoid the repeated simulation for the same solution and increase the execution efficiency. The approach advanced in this study may be used to improve greatly such complex maintenance decision making by taking into account the uncertainties. It can help planners to make decisions regarding solution feasibility, total maintenance possession time and system availability.

Acknowledgement

This study was done whilst Tao Zhang was visiting in the University of Nottingham. His research and visiting were supported by the National Science Foundation of China, under Grant No. 70971132 and 60904002. The authors gratefully acknowledge Prof. John Andrews who is the Royal Academy of Engineering and Network Rail Professor of Infrastructure Asset Management. He is also the Director of The Lloyd's Register Educational Trust Centre for Risk and Reliability Engineering at the University of Nottingham. Thank him for many helpful and constructive comments.

References

1. Cho D I, Parlar M. A survey of maintenance models for multi-unit systems. European Journal of Operational Research, 1991, 51:1–23.
2. Dekker R F, Schouten D, Wildeman R. A review of multi-component maintenance models with economical dependence. Mathematical Methods of Operations Research, 1996, 45:411–435.
3. Wang H. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research, 2002, 139:469–489.
4. Nicolai R P, Dekker R. Optimal Maintenance of Multi-Component Systems: a Review//K.A.H.Kobbacy, D.N.P.Murthy. Complex System Maintenance Handbook. 2006, London: Springer Verlag.
5. Sheu S H, Griffith W S. Extended block replacement policy with shock models and used Items. European Journal of Operational Research. 2002, 140: 50–60.
6. Castainer B, Grall A, Berenguer C. A condition-based maintenance policy with non-periodic inspections for a two-unit series system. Reliability Engineering and System Safety, 2005, 87:109–120.
7. Laggoune R, Alaa C, Djamil A. Impact of few failure data on the opportunistic replacement policy for multi-component systems. Reliability Engineering and System Safety, 2010, 95:108–119.
8. Salah O M, Daoud A K, Ali G. A simulation model for opportunistic maintenance strategies. the 1999 7th IEEE International Conference on Emerging Technologies and Factory Automation. 1999, 703–709.
9. Marseguerra M, Zio E, Podofillini L. Condition-based maintenance optimization by means of genetic algorithms and Monte Carlo simulation. Reliability Engineering and System Safety. 2002, 77:151–166.
10. Barata J, Soares C G, Marseguerra M, Zio E. Simulation modeling of repairable multi-component deteriorating system for on condition maintenance optimization. Reliability Engineering and System Safety. 2002, 76:255–264.

Tao ZHANG, Ph.D.
Zhijun CHENG, Ph.D.
Ya-jie LIU, Ph.D.
Prof. Bo GUO, Ph.D.
College of Information System and Management
National University of Defense Technology
47 Yanwachi Street, Changsha, China
E-mail: zhangtao@nudt.edu.cn, 197697@qq.com, boguo@nudt.edu.cn