Scientific Board

Professor Andrzej Niewczas, PhD, DSc (Eng)

Chair of Scientific Board
President of the Board of the Polish Maintenance Society

Professor Holm Altenbach, PhD, DSc (Eng)
Martin Luther Universität, Halle - Wittenberg, Germany

Professor Zdzisław Chłopek, PhD, DSc (Eng)
Warsaw University of Technology, Warsaw

Professor Jan Dąbrowski, PhD, DSc (Eng)
Białystok Technical University, Białystok

Professor Sławczo Denczew, PhD, DSc (Eng)
Warsaw University of Technology, School of Technical and Social Sciences in Płock

Dr Ilia Frenkel
Shamoon College of Engineering, Beer Sheva, Israel

Professor Olgierd Hryniewicz, PhD, DSc (Eng)
Systems Research Institute of the Polish Academy of Science, Warsaw

Professor Hong-Zhong Huang, PhD, DSc
University of Electronic Science and Technology of China, Chengdu, Sichuan, China

Professor Krzysztof Kołowrocki, PhD, DSc
Gdynia Maritime University

Professor Štefan Liščak
Žilinská univerzita, Žilina, Slovak Republic

Professor Vaclav Legat
Czech University of Agriculture, Prague, Czech Republic

Professor Leonas Povilas Lingaitis
Vilnius Gediminas Technical University, Vilnius, Lithuania

Professor Jerzy Merkisz, PhD, DSc (Eng)
Poznań University of Technology, Poznań

Professor Gilbert De Mey
University of Ghent, Belgium

Professor Tomasz Nowakowski, PhD, DSc (Eng)
Wrocław University of Technology, Wrocław

Professor Marek Orkisz, PhD, DSc (Eng)
Rzeszów University of Technology, Rzeszów

Professor Stanisław Piasecki, PhD, DSc (Eng)
Systems Research Institute of the Polish Academy of Sciences, Warsaw

Professor Leszek Piaseczny, PhD, DSc (Eng)
Polish Naval Academy, Gdynia

Professor Stanisław Radkowski, PhD, DSc (Eng)
Warsaw University of Technology, Warsaw

Professor Andrzej Seweryn, PhD, DSc (Eng)
Białystok Technical University, Białystok

Professor Zbigniew Smalko, PhD, DSc (Eng)
Air Force Institute of Technology, Warsaw

Professor Marian Szczerek, PhD, DSc (Eng)
Institute for Sustainable Technologies, Radom

Professor Jan Szybka, PhD, DSc (Eng)
AGH University of Science and Technology, Cracow

Professor Katsumi Tanaka
Kyoto University, Kyoto, Japan

Professor David Vališ, PhD, DSc (Eng)
University of Defence, Brno, Czech Republic

Professor Irina Yatskiv
Riga Transport and Telecommunication Institute, Latvia

Co-financed by the Minister of Science and Higher Education

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

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

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

Our IF is 0.333

Editorial staff:	Dariusz Mazurkiewicz, PhD, DSc (Eng), Associate Professor (Editor-in-Chief, Secretary of the Scientific Board) Tomasz Klepka, PhD (Eng) (Deputy Editor-in-Chief) Teresa Błachnio-Krolopp, MSc (Eng) (Editorial secretary) Andrzej Koma (Typesetting and text makeup) Krzysztof Olszewski, PhD (Eng) (Webmaster)
Publisher:	Polish Maintenance Society, Warsaw
Scientific patronage:	Polish Academy of Sciences Branch in Lublin
Address for correspondence:	"Eksploatacja i Niezawodność" - Editorial Office 20-618 Lublin, ul. Nadbystrzycka 36, Poland e-mail: office@ein.org.pl http://www.ein.org.pl/
Circulation:	550 copies

INFORMATION FOR AUTHORS

Terms and Conditions of Publication:

- The quarterly „Maintenance and Reliability" publishes original papers written in Polish with an English translation.
- Translation into English is done by the Authors after they have received information from the Editorial Office about the outcome of the review process and have introduced the necessary modifications in accordance with the suggestions of the referees!
- Acceptance of papers for publication is based on two independent reviews commissioned by the Editor.

Fees:

- Pursuant to a resolution of the Board of PNTTE, as of 2009 the publication fee for one text is 600 zloty + VAT.
- Coloured graphical elements in the submitted text require agreement from the Editor and are charged extra.

Technical requirements:

- After receiving positive reviews and after acceptance of the paper for publication, the text must be submitted in a Microsoft Word document format.
- Drawings and photos should be additionally submitted in the form of graphical files in the *.tif, *.jpg or *.cdr (v. X3) formats.
- A manuscript should include (in accordance with the enclosed correct manuscript format: *.pdf, *.doc):
- names of authors, title, abstract, and key words that should complement the title and abstract (in Polish and in English)
- the text in Polish and in English with a clear division into sections (please, do not divide words in the text);
- tables, drawings, graphs, and photos included in the text should have descriptive two-language captions,
- if this can be avoided, no formulae and symbols should be inserted into text paragraphs by means of a formula editor
- references (written in accordance with the required reference format)
- author data – first names and surnames along with scientific titles, affiliation, address, phone number, fax, and e-mail address
- The Editor reserves the right to abridge and adjust the manuscripts.
- All submissions should be accompanied by a submission form.

Editor contact info: (Submissions should be sent to the Editor's address)

Editorial Office of „Eksploatacja i Niezawodność - Maintenance and Reliability"
Nadbystrzycka 36, 20-618 Lublin, Poland
e-mail: office@ein.org.pl

INFORMATION FOR SUBSCRIBERS

Fees

Yearly subscription fee (four issues) is 100 zloty and includes delivery costs.
Subscribers receive any additional special issues published during their year of subscription free of charge.

Orders

Subscription orders along with authorization to issue a VAT invoice without receiver's signature should be sent to the Editor's address.

Note

In accordance with the requirements of citation databases, proper citation of publications appearing in our Quarterly should include the full name of the journal in Polish and English without Polish diacritical marks, i.e.,
Eksploatacja i Niezawodnosc – Maintenance and Reliability.

No text or photograph published in „Maintenance and Reliability" can be reproduced without the Editor's written consent.

Science and Technology

KOMOREK A, PRZYBYŁEK P. **Examination of the influence of cross-impact load on bend strength properties of composite materials, used in aviation**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 265–269.

Fibre reinforced composites are often used in airplane structures because of their specific strength. One type of the materials are layered composites (laminates) applied inter alia in aircraft's covering production. Laminate is susceptible to damage resulting from impacts, the effect of which is usually invisible during macroscopic observation. The article presents results of a preliminary examination of layered composites obtained from an airplane element loaded impactly with low energy. During testing, pieces were loaded with 2.5; 5 and 10 J energy and then they were put on bend tests. The material bending strength after a shock load with 2.5 and 5 J energy remains almost unaltered, but for 10 J energy, it decreases by more than 30% in comparison to undamaged material. As a result of the examination, it was ascertained that in all cases the exact location of the damage could be difficult to find, which is a significant maintenance problem.

JIANG G, ZHU M, WU Z. **Reliability allocation using probabilistic analytical target cascading with efficient uncertainty propagation**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 270–277.

Analytical target cascading (ATC) provides a systematic approach in solving reliability allocation problems for large scale system consisting of a large number of subsystems, modules and components. However, variability and uncertainty in design variables (e.g., component reliability) are usually inevitable, and when they are taken into consideration, the multi-level optimization will be very complicated. The impacts of uncertainty on system reliability are considered in this paper within the context of probabilistic ATC (PATC) formulation. The challenge is to reformulate constraints probabilistically and estimate uncertainty propagation throughout the hierarchy since outputs of subsystems at lower levels constitute inputs of subsystems at higher levels. The performance measure approach (PMA) and the performance moment integration (PMI) method are used to deal with the two objectives respectively. To accelerate the probabilistic optimization in each subsystem, a unified framework for integrating reliability analysis and moment estimation is proposed by incorporating PATC with single-loop method. It converts the probabilistic optimization problem into an equivalent deterministic optimization problem. The computational efficiency is remarkably improved as the lack of iterative process during uncertainty analysis. A nonlinear geometric programming example and a reliability allocation example are used to demonstrate the efficiency and accuracy of the proposed method.

KUMAR J, KADYAN M S, MALIK S C. **Cost analysis of a two-unit cold standby system subject to degradation, inspection and priority**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 278–283.

The present paper deals with a reliability model incorporating the idea of degradation, inspection and priority. The units may fail completely directly from normal mode. There is a single server who visits the system immediately when required. The original unit undergoes for repair upon failure while only replacement of the duplicate unit is made by similar new one. The original unit does not work as new after repair and so called degraded unit. The system is considered in up-state if any one of new/duplicate/degraded unit is operative. The server inspects the degraded unit at its failure to see the feasibility of repair. If repair of the degraded unit is not feasible, it is replaced by new one similar to the original unit in negligible time. The priority for operation to the new unit is given over the duplicate unit. The distribution of failure time follow negative exponential where as the distributions of inspection, repair and replacement times are assumed as arbitrary. The system is observed at suitable regenerative epochs by using regenerative point technique to evaluate mean time to system failure (MTSF), steady-state availability, busy period and expected number of visits by the server. A particular case is considered to see graphically the trend of mean time to system failure (MTSF), availability and profit with respect to different parameters.

HUANG HZ. **Structural reliability analysis using fuzzy sets theory**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 284–294.

Prediction of structural performance is a complex problem because of the existence of randomness and fuzziness in engineering practice. In this area, reliability analyses have been performed using probabilistic methods. This work investigates reliability analysis of structure involving fuzziness and randomness. In particular, the safety state of the structure is defined by a fuzzy state variable, fuzzy random allowable interval, or fuzzy random generalized strength. Because the membership function of the fuzzy safety state is the key to structural reliability analysis using the fuzzy sets theory, this work proposes useful methods to determine the membership functions and develops a

KOMOREK A, PRZYBYŁEK P. **Badanie wpływu poprzecznych obciążeń udarowychna właściwości wytrzymałościowe materiałów kompozytowych stosowanych w lotnictwie**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 265–269.

Kompozyty włókniste ze względu na bardzo wysoką wytrzymałość właściwą są często stosowane w konstrukcjach lotniczych. Jedną z odmian tych materiałów są kompozyty warstwowe (laminaty), z których wykonuje się m.in. elementy pokryć statków powietrznych. Laminat jest materiałem wrażliwym na działanie porzecznych obciążeń udarowych, często występujących podczas eksploatacji samolotów i śmigłowców. Praca prezentuje wyniki wstępnych badań kompozytów warstwowych pobranych z rzeczywistego elementu lotniczego, poddanych niskoenergetycznym obciążeniom udarowym. Podczas eksperymentu, próbki obciążano energiami o wartościach 2,5; 5 i 10 J, a następnie poddawano próbom zginania. Wytrzymałość materiału po obciążeniu udarowym z energiami 2,5 i 5 J pozostaje niemal niezmieniona, natomiast dla energii 10 J spada o ponad 30% w stosunku do materiału nieuszkodzonego. W wyniku badań stwierdzono również, że w każdym z przypadków mogą wystąpić trudności z lokalizacją uszkodzenia, co stanowi istotny problem eksploatacyjny.

JIANG G, ZHU M, WU Z. **Alokacja niezawodności z wykorzystaniem probabilistycznej metody analitycznego kaskadowania celówzapewniająca wydajną propagację niepewności**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 270–277.

Analityczne kaskadowanie celów (ATC) stanowi systematyczne podejście do rozwiązywania zagadnień alokacji niezawodności dotyczących systemów wielkoskalowych składających się z dużej liczby podsystemów, modułów i elementów składowych. Jednakże zmienność i niepewność zmiennych projektowych (np. niezawodności elementów składowych) są zazwyczaj nieuniknione, a gdy weźmie się je pod uwagę, optymalizacja wielopoziomowa staje się bardzo skomplikowana. W prezentowanym artykule, wpływ niepewności na niezawodność systemu rozważano w kontekście formuły probabilistycznego ATC (PATC). Wyzwanie polegało na probabilistycznym przeformułowaniu ograniczeń oraz ocenie propagacji niepewności w całej hierarchii, jako że wyjścia podsystemów na niższych poziomach stanowią wejścia podsystemów na poziomach wyższych. Cele te realizowano, odpowiednio, przy użyciu metody minimum funkcji granicznej (performance measure approach, PMA) oraz metody całkowania momentów statystycznych funkcji granicznej (performance moment integration, PMI). W celu przyspieszenia probabilistycznej optymalizacji w każdym podsystemie, zaproponowano ujednolicone ramy pozwalające na integrację analizy niezawodności z oceną momentów statystycznych poprzez połączenie PATC z metodą jednopoziomową (pojedynczej pętli, single-loop method). Zaproponowana metoda polega na przekształceniu probabilistycznego zagadnienia optymalizacyjnego na deterministyczne zagadnienie optymalizacyjne. Zwiększa to znacznie wydajność obliczeniową w związku z brakiem procesu iteratywnego podczas analizy niepewności. Wydajność i trafność proponowanej metody wykazano na podstawie przykładów dotyczących programowania nieliniowego geometrycznego oraz alokacji niezawodności.

KUMAR J, KADYAN M S, MALIK S C. **Analiza kosztów dwu-elementowego systemu z rezerwą zimną z uwzględnieniem degradacji, kontroli stanu systemu oraz priorytetowości zadań.** Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 278–283.

Niniejsza praca dotyczy modelu niezawodności uwzględniającego zagadnienia degradacji, kontroli stanu oraz priorytetowości zadań. Elementy mogą ulegać całkowitemu uszkodzeniu bezpośrednio z trybu normalnego. Istnieje jeden konserwator, który odwiedza system, gdy tylko zachodzi taka potrzeba. W przypadku uszkodzenia, element oryginalny podlega naprawie, podczas gdy element zapasowy (duplikat) podlega jedynie wymianie na nowy, podobny. Po naprawie, element oryginalny nie działa już jako element nowy lecz jako element zdegradowany. System uważa się za zdatny jeżeli pracuje którykolwiek z trzech typów elementów: nowy/rezerwowy/zdegradowany. W przypadku uszkodzenia elementu zdegradowanego, konserwator przeprowadza kontrolę stanu elementu, aby stwierdzić możliwość realizacji naprawy. Jeżeli naprawa elementu zdegradowanego jest niemożliwa, zostaje on wymieniony, w czasie pomijalnym, na element nowy, podobny do elementu oryginalnego. Nowy element uzyskuje priorytet pracy w stosunku do elementu rezerwowego. Rozkład czasu uszkodzenia jest rozkładem wykładniczym ujemnym, a rozkłady czasów kontroli stanu, naprawy i wymiany przyjmuje się jako rozkłady dowolne. System obserwuje się w odpowiednich okresach odnowy wykorzystując technikę odnowy RPT (regenerative point technique) w celu ocenienia średniego czasu do uszkodzenia systemu (MTSF), gotowości stacjonarnej, okresu zajętości oraz oczekiwanej liczby wizyt konserwatora. Przebiegi MTSF, gotowości i zysków w funkcji różnych parametrów przedstawiono w formie graficznej na podstawie studium przypadku.

HUANG HZ. **Analiza niezawodnościowa konstrukcji z wykorzystaniem teorii zbiorów rozmytych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 284–294.

Przewidywanie zachowania konstrukcji stanowi złożone zagadnienie ze względu na istnienie w praktyce inżynierskiej losowości i rozmytości. Na tym obszarze, analizy niezawodnościowe prowadzono dotąd przy pomocy metod probabilistycznych. W niniejszej pracy przedstawiono metodę niezawodnościowej analizy konstrukcji uwzględniającą rozmytość i losowość. Dokładniej, stan bezpieczeństwa konstrukcji określano za pomocą rozmytej zmiennej stanu, rozmytego losowego przedziału dozwolonego lub rozmytej losowej uogólnionej wytrzymałości. Ponieważ funkcja przynależności rozmytego stanu structural reliability analysis method based on the fuzzy safety state. Several examples are provided to illustrate the proposed methods.

bezpieczeństwa stanowi klucz do niezawodnościowej analizy konstrukcji wykorzystującej teorię zbiorów rozmytych, w niniejszej pracy zaproponowano przydatne metody wyznaczania funkcji przynależności oraz opracowano metodę niezawodnościowej analizy konstrukcji opartą na rozmytym stanie bezpieczeństwa. Zaproponowane metody zilustrowano kilkoma przykładami.

JURECKI R, JAŚKIEWICZ M, GUZEK M, LOZIA Z, ZDANOWICZ P. **Driver's reaction time under emergency braking a car – research in a driving simulatory**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 295–301.

Extract. The paper presents results of research on the reaction time of drivers of motor vehicles in case of accident risk. These studies have been conducted in an autoPW driving simulator - research project N509 016 31/1251. During a simulated pre-accident situation, a test driver, as he was trying to avoid a collision with an obstacle, has been forced to perform a braking manoeuvre. 107 people of different age and driving experience have been studied. Various scenarios describing risk situations (the speed of a car driven by a tested driver, the distance between the vehicles at the initial moment) have been considered.Reaction times were determined at the controls of the vehicle and they have been presented in a time to collision function. The presented results provide important information on the analysed subject (for instance reaction time dependence on risk time). They have also confirmed the usefulness of this type of simulation studies in connection with the possibility of fairly accurate reproduction of the environment appearance and an event scenario in relation to possible real situations.

JURECKI R, JAŚKIEWICZ M, GUZEK M, LOZIA Z, ZDANOWICZ P. **Czas reakcji kierowcy w warunkach awaryjnego hamowania samochodu – badania w symulatorze jazdy samochodem**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 295–301.

W publikacji przedstawiono wyniki badań dotyczące czasu reakcji kierowców pojazdów samochodowych w sytuacjach zagrożenia wypadkowego. Badania te zostały wykonane w symulatorze jazdy samochodem autoPW w ramach projektu badawczego N509 016 31/1251. W zainscenizowanej sytuacji przedwypadkowej, badany kierowca próbując uniknąć zderzenia z przeszkodą, zmuszony był do wykonania manewru hamowania. Przebadano 107 osób różniących się wiekiem i doświadczeniem w prowadzeniu samochodu. Rozważono różne scenariusze opisujące sytuacje zagrożenia (prędkość samochodu, którym kierował badany, odległość pomiędzy pojazdami w chwili początkowej). Wyznaczono czasy reakcji na elementach sterowania pojazdem i przedstawiono je w funkcji czasu ryzyka (ang. time to collision). Zaprezentowane wyniki dostarczyły istotnych informacji na analizowany temat (np. zależność czasu reakcji od czasu ryzyka). Potwierdziły też przydatność prowadzenia tego rodzaju badań w symulatorach w związku z możliwością dosyć wiernego odwzorowania wyglądu otoczenia i scenariusza zdarzeń w stosunku do możliwych rzeczywistych sytuacji.

ZHANG Z, GAO W, ZHOU Y, ZHANG Z. **Reliability modeling and maintenance optimization of the diesel system in locomotives**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 302–311.

Engine system is a prone-fault part in diesel locomotive and its malfunctions always occur regularly in different seasons in practice. However, the current maintenance policy in China has not attached deserving importance to seasonal influence, which is considered as one of the main causes for over/under-maintenance. To assess the current maintenance, in this study a double-fold Weibull competing risk model for summer and winter is developed using the real failure data (2008-2011) of locomotives from Urumqi Railway Bureau. Meanwhile, a new approach, termed as Approximately Combined Parameter Method (ACPM), is proposed to combine the initially estimated parameters into different folds, which can avoid a subjective determination of the model's parameters fold. After that, the combined parameters are used as initial values for maximum likelihood estimate (MLE) to achieve an accurate model. Necessary optimizations are introduced based on the chosen models. Results show that the maintenance period differs a lot between winter and summer, and the optimized maintenance can increase the availability and decrease cost more than the existing policy

ZHANG Z, GAO W, ZHOU Y, ZHANG Z. **Modelowanie niezawodności i optymalizacja utrzymania ruchu układu samoczynnego zapłonu w lokomotywach**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 302–311.

Układ silnikowy stanowi podatną na uszkodzenia część lokomotywy spalinowej, a w praktyce jego awarie występują zawsze regularnie w zależności od pory roku. Pomimo tego, obecna polityka obsługowa w Chinach nie przywiązuje wystarczającej wagi do wpływu pór roku, co uważa się za główną przyczynę nadmiernych lub niedostatecznych działań obsługowych. Aby ocenić bieżące działania obsługowe, w niniejszym artykule opracowano model zagrożeń konkurujących dla lata i zimy, oparty na połączeniu dwóch rozkładów Weibulla, wykorzystujący rzeczywiste dane o uszkodzeniach (2009–2011) lokomotyw używanych przez Agencję Kolejową Urumqui. Jednocześnie zaproponowano nowe podejście, o nazwie Approximately Combined Parameter Method (Metoda Przybliżonego Łączenia Parametrów, ACPM), które polega na łączeniu wstępnie obliczonych parametrów w różne wielokrotności, co pozwala na uniknięcie subiektywnego wyznaczania liczby parametrów modelu. W celu otrzymania dokładnego modelu, połączone parametry wykorzystuje się jako wstępne wartości w estymacji metodą największej wiarygodności. Konieczne optymalizacje wprowadza się na podstawie wybranych modeli. Wyniki pokazują, że letni okres obsługowy różni się zasadniczo od zimowego, a zoptymalizowana obsługa może zwiększyć gotowość systemu i zmniejszyć koszty utrzymania ruchu w większym stopniu niż dotychczasowa polityka obsługowa.

LIU Y, HUANG HZ, WANG Z, LI YF, ZHANG XL. **Joint optimization of redundancy and maintenance staff allocation for multi-state series-parallel systems**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 312–318.

Multi-state system (MSS), as a kind of complex system consisting of elements with different performance levels, widely exists in engineering practices. In this paper, redundancy and maintenance staff allocation problems for repairable MSS with series-parallel configuration are considered simultaneously. The traditional redundancy allocation problem (RAP) for MSS always assumes that maintenance resources are unlimited. However in many practical situations, maintenance resources are limited due to the budget and/or time. To maximize the system availability under a certain demand, there are two feasible ways: (1) designing an optimal system configuration with available elements, and (2) allocating more maintenance staffs to reduce waiting time for repair. With the assistance of Markov queue model, the availabilities of identical version elements with the pre-assigned number of maintenance staffs can be evaluated. The universal generation function (UGF) is employed to assess the availability of entire MSS under a certain demand. Two optimization formulas considering the limited maintenance resources are proposed. One regards the limitation of maintenance resources as a constraint, and the other considers minimizing the total system cost including both the system elements and maintenance staff fees. The system redundancy and staffs allocation strategies are jointly optimized under required availability. A numerical case is presented to illustrate the efficiency of the proposed models. The Firefly Algorithm (FA), which is a recently developed metaheuristic optimization algorithm, is employed to seek the global optimal strategy.

LIU Y, HUANG HZ, WANG Z, LI YF, ZHANG XL. **Optymalizacja łączona alokacji nadmiarowości oraz alokacji pracowników służb utrzymania ruchu w wielostanowych systemach szeregowo-równoległych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 312–318.

Systemy wielostanowe (multi-state systems, MSS), stanowiące typ złożonych systemów zbudowanych z elementów o różnym poziomie wydajności, znajdują szerokie zastosowanie w praktyce inżynierskiej. W prezentowanej pracy podjęto rozważania łączące zagadnienia alokacji nadmiarowości oraz alokacji pracowników służb utrzymania ruchu w naprawialnych systemach MSS o konfiguracji szeregowo-równoległej. Tradycyjnie ujmowane zagadnienie alokacji nadmiarowości (redundancy allocation problem, RAP) w systemach MSS zawsze zakłada, że środki obsługi są nieograniczone. Jednakże w wielu sytuacjach praktycznych, środki obsługi mogą być ograniczone budżetem i/lub czasem. Istnieją dwa możliwe sposoby maksymalizacji gotowości systemu przy określonym zapotrzebowaniu użytkowników: (1) zaprojektowanie optymalnej konfiguracji systemu z wykorzystaniem dostępnych elementów oraz (2) alokowanie większej liczby pracowników obsługi w celu zmniejszenia czasu oczekiwania na naprawę. Dostępność jednakowych wersji elementów przy wcześniej określonej liczbie pracowników obsługi oceniano za pomocą modelu kolejek Markowa. Uniwersalną funkcję generacyjną (UGF) wykorzystano do oceny gotowości całego systemu MSS przy określonym zapotrzebowaniu. Zaproponowano dwa równania optymalizacyjne uwzględniające ograniczone środki obsługi. W jednym z nich ograniczoność środków obsługi potraktowano jako ograniczenie (constraint), natomiast drugie równanie dotyczyło minimalizacji całkowitych kosztów systemu włącznie z kosztami elementów systemu oraz płacą pracowników służb utrzymania ruchu. Strategie alokacji nadmiarowości systemu oraz alokacji pracowników poddano jednoczesnej optymalizacji z uwzględnieniem wymaganej gotowości. Wydajność proponowanych modeli zilustrowano przykładem numerycznym. Poszukiwania optymalnej strategii globalnej prowadzono przy pomocy niedawno opracowanego metaheurystycznego algorytmu optymalizacyjnego znanego jako algorytm świetlika (Firefly Algorithm, FA).

KAMIŃSKI Z, CZABAN J. **Diagnosing of the agricultural tractor braking system within approval tests**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 319–326.

The paper describes the requirements for type approval testing of agricultural tractors with regard to braking, while taking into account the draft Regulation drawn up by the European Commission's Working Group on Agricultural Tractors (WGAT). A programme and methodology for testing the performance of the tractor brake system and its air brake system to supply and control the braking system of a towed vehicle are presented. Examples of diagnostic tests of a prototype tractor with hydraulically actuated brakes are given. These diagnostic tests may be wholly or partially used to develop a programme of qualification testing of tractors on production lines and a programme of periodic technical inspections of operated tractors.

TEKCAN A T, KAHRAMANOĞLU G, GÜNDÜZALP M. **Incorporating product robustness level in field return rate predictions**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 327–332.

Reliability and return rate prediction of products are traditionally achieved by using stress based standards and/or applying accelerated life tests. But frequently, predicted reliability and return rate values by using these methods differ from the field values. The primary reason for this is that products do not only fail due to the stress factors mentioned in the standards and/or used in accelerated life tests. There are additional failure factors, such as ESD, thermal shocks, voltage dips, interruptions and variations, quality factors, etc. These factors should also be considered in some way when predictions are made during the R&D phase. Therefore, a method should be used which considers such factors, thus increasing the accuracy of the reliability and return rate prediction. In this paper, we developed a parameter, which we call Robustness Level Factor, to incorporate such factors, and then we combined this parameter with traditional reliability prediction methods. Specifically, the approach takes into account qualitative reliability tests performed during the R&D stage and combines them with life tests by using Artificial Neural Networks (ANN). As a result, the approach gives more accurate predictions compared with traditional prediction methods. With this prediction model, we believe that analysts can determine the reliability and return rate of their products more accurately.

BUGARIC U, TANASIJEVIC M, POLOVINA D, IGNJATOVIC D, JOVANCIC P. **Lost production costs of the overburden excavation system caused by rubber belt failure**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 333–341.

In this paper the average malfunction costs (lost production) of the overburden excavation system on the Tamnava - East field open-pit mine caused by the failure of belt conveyor rubber belts which work on the bucket wheel excavator, belt wagon and spreader are determined, i.e. the unit cost of system malfunction per hour of belt conveyor work during belt lifetime. The basis for the calculation of malfunction costs is presented by the proposed methodology for the analysis of rubber belt working time to failure based on the fact that working time until sudden failure (tear, breakthrough) can be described by exponential distribution while working time until gradual failure can be described by normal distribution. The proposed methodology as well as the expression for malfunction cost determination can be used, with appropriate adoptions, in the analysis of the functioning of other open-pit mines for better planning of malfunctions, requirements for spare rubber belts as well reductions in working costs, i.e. they can indicate better (optimal) maintenance strategy.

SU C, ZHANG YJ, CAO BX. **Forecast model for real time reliability of storage system based on periodic inspection and maintenance data**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 342–348.

In recent years, storage reliability has attracted much attention for increasing reliability requirement. In this paper, forecast models for real-time reliability of storage system under periodic inspection and maintenance are presented, which is based on the theories of reliability physics and exponential distribution. The models are developed under two newly-defined imperfect repair modes, i.e., Improved As Bad As Old (I-ABAO) and Improved As Good As New (I-AGAN). A completion method for censored life data is also proposed by averaging the residual lifetime. According to the complete and censored lifetime data, parameters in the models are estimated by applying maximum likelihood estimation method and iterative method respectively. A numerical example of a storage system is given to verify the feasibility of the proposed completion method and the effectiveness of the two models.

KAMIŃSKI Z, CZABAN J. **Diagnostyka układu hamulcowego ciągnika rolniczegow ramach badań homologacyjnych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 319–326.

W pracy opisano wymagania dotyczące badań homologacyjnych ciągników rolniczych w zakresie hamowania, uwzględniające propozycje przepisów opracowywanych przez grupę roboczą Komisji Europejskiej ds. ciągników rolniczych (Working Group on Agricultural Tractors - WGAT). Przedstawiono program i metodykę badań skuteczności układu hamulcowego ciągnika oraz jego instalacji powietrznej do zasilania i sterowania hamulcami pojazdu ciągniętego. Zamieszczono przykłady testów diagnostycznych prototypowego ciągnika z hamulcami uruchamianymi hydraulicznie. Opisane testy diagnostyczne mogą być w całości lub częściowo wykorzystane do opracowania programu badań kwalifikacyjnych ciągników na liniach produkcyjnych oraz programu okresowych badań technicznych ciągników eksploatowanych.

TEKCAN A T, KAHRAMANOĞLU G, GÜNDÜZALP M. **Przewidywanie rzeczywistego wskaźnika zwrotów towaruz uwzględnieniem poziomu odporności produktu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 327–332.

Niezawodność i wskaźniki zwrotów towaru przewiduje się tradycyjnie przy użyciu norm obciążeniowych i/lub stosując przyspieszone badania trwałości. Jednakże, często wartości niezawodności i wskaźnika zwrotów przewidywane za pomocą tych metod różnią się od ich wartości rzeczywistych. Główną tego przyczyną jest fakt, że produkty nie ulegają awarii wyłącznie pod wpływem czynników obciążeniowych wymienianych w normach i/lub wykorzystywanych w przyspieszonych badaniach trwałości. Istnieją dodatkowe czynniki wpływające na intensywność uszkodzeń, takie jak wyładowania elektrostatyczne, wstrząsy termiczne, spadki, przerwy w dostawie i zmiany napięcia, czynniki jakościowe, itp. Te czynniki także powinny być w jakiś sposób uwzględnione przy dokonywaniu predykcji na etapie badań i rozwoju (R&D). Dlatego też zwiększenie trafności predykcji niezawodności i wskaźników zwrotów towaru wymaga metody, która uwzględniałaby tego typu czynniki. W niniejszej pracy opracowaliśmy parametr, nazwany przez nas "czynnikiem poziomu odporności", który pozwala na uwzględnienie takich czynników, a następnie wykorzystaliśmy ów parametr w połączeniu z tradycyjnymi metodami przewidywania niezawodności. W szczególności, przedstawione podejście bierze pod uwagę jakościowe badania niezawodnościowe wykonywane na etapie R&D łącząc je z badaniami trwałościowymi przy użyciu sztucznych sieci neuronowych ANN. Dzięki temu, w podejściu tym uzyskuje się bardziej trafne predykcje niż w tradycyjnych metodach prognozowania. Jesteśmy przekonani, że użycie powyższego modelu predykcyjnego umożliwi analitykom bardziej trafne wyznaczanie niezawodności oraz wskaźników zwrotów wytwarzanych przez nich produktów.

BUGARIC U, TANASIJEVIC M, POLOVINA D, IGNJATOVIC D, JOVANCIC P. **Koszty utraconej produkcji wywołane uszkodzeniem gumowej taśmy transportowej w układzie maszynowym do zdejmowania nadkładu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 333–341.

W prezentowanym artykule wyznaczono średnie koszty awarii (utraconej produkcji) układu maszynowego do zdejmowania nadkładu, wykorzystywanego w kopalni odkrywkowej Tamnava - East field, spowodowanej uszkodzeniem przenośnikowych taśm gumowych zastosowanych w koparce wielonaczyniowej, samobieżnym przenośniku taśmowym oraz zwałowarce. Koszt zdefiniowano jako jednostkowy koszt awarii układu na godzinę pracy przenośnika taśmowego podczas cyklu życia taśmy. Podstawę obliczeń kosztów awarii stanowiła zaproponowana metoda analizy czasu pracy taśmy gumowej do uszkodzenia, oparta na fakcie, iż czas pracy do nagłego uszkodzenia (rozerwanie, przebicie) można opisać za pomocą rozkładu wykładniczego, natomiast czas pracy do stopniowego uszkodzenia – za pomocą rozkładu normalnego. Proponowana metodologia, jak również równania do wyznaczania kosztów awarii mogą być wykorzystywane, przy odpowiedniej adaptacji, do analizy funkcjonowania innych kopalni odkrywkowych służąc lepszemu planowaniu awarii, zapotrzebowania na zapasowe pasy gumowe oraz redukcji kosztów pracy. Mogą one, innymi słowy, wskazywać (optymalną) strategię utrzymania ruchu.

SU C, ZHANG YJ, CAO BX. **Model do prognozowania niezawodności systemu magazynowania w czasie rzeczywistym w oparciu o dane z przeglądów okresowych oraz dane eksploatacyjne**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 342–348.

W ostatnich latach wiele uwagi poświęcono tematyce niezawodności magazynowania w odniesieniu do zwiększania wymogu niezawodności. W prezentowanym artykule, przedstawiono modele do prognozowania w czasie rzeczywistym niezawodności systemu magazynowania podlegającego przeglądom okresowym i obsłudze. Modele oparto na teoriach z zakresu fizyki niezawodności oraz na rozkładzie wykładniczym. Proponowane modele opracowano dla dwóch nowo zdefiniowanych opcji niepełnej odnowy, t.j. Improved-As Bad As Old (Jak Tuż Przed Uszkodzeniem – Wersja Udoskonalona) oraz Improved-As Good As New (Jak Fabrycznie Nowy – Wersja Udoskonalona). Zaproponowano także metodę uzupełniania danych cenzurowanych (uciętych) dotyczących trwałości polegającą na uśrednianiu trwałości resztkowej. Zgodnie z pełnymi i cenzurowanymi danymi trwałościowymi, parametry w proponowanych modelach ocenia się, odpowiednio, z zastosowaniem estymacji metodą największej wiarygodności oraz metody iteracyjnej. Poprawność przedstawionej metody uzupełniania oraz efektywność proponowanych dwóch modeli zweryfikowano na podstawie numerycznego przykładu systemu magazynowania.

Article citation info:
KOMOREK A, PRZYBYŁEK P. Examination of the influence of cross-impact load on bend strength properties of composite materials, used in aviation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 265–269.

Andrzej KOMOREK
Paweł PRZYBYŁEK

EXAMINATION OF THE INFLUENCE OF CROSS-IMPACT LOAD ON BEND STRENGTH PROPERTIES OF COMPOSITE MATERIALS, USED IN AVIATION

BADANIE WPŁYWU POPRZECZNYCH OBCIĄŻEŃ UDAROWYCH NA WŁAŚCIWOŚCI WYTRZYMAŁOŚCIOWE MATERIAŁÓW KOMPOZYTOWYCH STOSOWANYCH W LOTNICTWIE*

Fibre reinforced composites are often used in airplane structures because of their specific strength. One type of the materials are layered composites (laminates) applied inter alia in aircraft's covering production. Laminate is susceptible to damage resulting from impacts, the effect of which is usually invisible during macroscopic observation. The article presents results of a preliminary examination of layered composites obtained from an airplane element loaded impactly with low energy. During testing, pieces were loaded with 2.5; 5 and 10 J energy and then they were put on bend tests. The material bending strength after a shock load with 2.5 and 5 J energy remains almost unaltered, but for 10 J energy, it decreases by more than 30% in comparison to undamaged material. As a result of the examination, it was ascertained that in all cases the exact location of the damage could be difficult to find, which is a significant maintenance problem.

***Keywords**: composite, laminate, shock load, delamination, airplane.*

Kompozyty włókniste ze względu na bardzo wysoką wytrzymałość właściwą są często stosowane w konstrukcjach lotniczych. Jedną z odmian tych materiałów są kompozyty warstwowe (laminaty), z których wykonuje się m.in. elementy pokryć statków powietrznych. Laminat jest materiałem wrażliwym na działanie porzecznych obciążeń udarowych, często występujących podczas eksploatacji samolotów i śmigłowców. Praca prezentuje wyniki wstępnych badań kompozytów warstwowych pobranych z rzeczywistego elementu lotniczego, poddanych niskoenergetycznym obciążeniom udarowym. Podczas eksperymentu, próbki obciążano energiami o wartościach 2,5; 5 i 10 J, a następnie poddawano próbom zginania. Wytrzymałość materiału po obciążeniu udarowym z energiami 2,5 i 5 J pozostaje niemal niezmieniona, natomiast dla energii 10 J spada o ponad 30% w stosunku do materiału nieuszkodzonego. W wyniku badań stwierdzono również, że w każdym z przypadków mogą wystąpić trudności z lokalizacją uszkodzenia, co stanowi istotny problem eksploatacyjny.

***Słowa kluczowe**: kompozyt, laminat, obciążenie udarowe, delaminacja, statek powietrzny.*

1. Introduction

Fibre reinforced composites, due to their high specific strength (Rm/p) and other beneficial effects of mechanical properties are increasingly used in basic airplane structures. Laminates (layered composites) are often selected for airplane structures and components, whose minimum mass remains the main criterion. Although there are a number of unquestionable advantages in the process of design, production and particularly exploitation of composite structures, work conditions seem to play a much higher role. One should also bear in mind property differences between composites and other construction materials [10].

Metallic materials after applying a shock load become damaged in a way relatively easy to locate. However, in case of fibre reinforced composites, the damage seems to be extremely difficult to spot due to the fact that it can emerge on the surface directly on the side of the shock load, on the opposite side of the shock load or inside the material's structure (laminates) [3, 7]. In aviation, even minor damage to the composite structure may drastically lower the bottom margin of the structure's safety, even after an impact of a dropped tool, foreign bodies thrust by aircraft wheels during a take-off, due to mid-air bird strike, etc. This type of bending strength is characterized by applying a relatively small mass at low velocity, therefore the energy of the impact also reaches low values. It must be emphasized that in literature the term itself and assigning the bend strength after an impact are referred to as low-energy, high-energy, low-velocity, high-velocity impact, etc. [1, 14, 17].

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

2. Damage to fibre composites

Fibre reinforced composites (FRP – fibre reinforced polymer) are materials of heterogeneous anisotropic structure, which makes the damage uneasy to locate. The damage can appear in various forms [6]:

a) base damage – cracking of the base along fibres, caused by their extending, gripping or cutting (Fig. 1);
b) delamination (separation of layers) – caused by stresses among various layers of the laminate (Fig. 1);
c) fibre damage – it commonly appears on the surface directly under the impact of a shock load, as a result of cutting forces which arise while bending the acting element; also on the side of the laminate away from the bending tensions;
d) perforation – it is macroscopic damage which emerges after destroying the base and the fibres, finally leading to the puncture (destruction) of the material's structure.

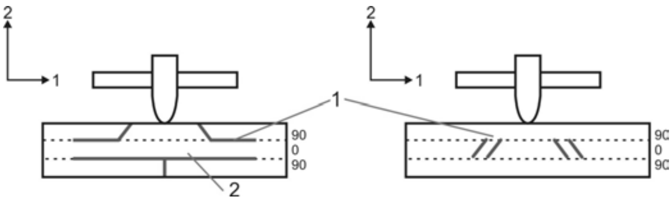


Fig. 1. Initial damage in 0/90/0 composite due to cross-impact loads: 1 – base damage, 2 – delamination [11]

The caused effects fall into two energy ranges of the impact load:

- low-energy damage (BVID – barely visible impact damage) as a result of small energy impact load. Within the material, around the impact area, there is a network of separated layers and cross-bending of the layers. However, there is no massive cracking of the fibres. On the surface of the damage there appears a slight spot, whereas on the opposite side the damage is much more extensive [2].
- high-energy damage with fibre cracking, partial or full perforation.

In case of minor impact loading (several joules) around the impact area, there is a slight spot, difficult to identify. The examination of the microstructure shows that beneath the spot there may emerge an area similar to a cut cone (Fig. 2). The tip of the cone overlaps the impacted spot (external surface of the element), whereas the base of the cone

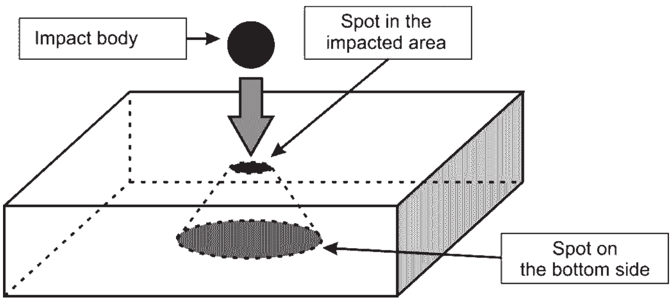


Fig. 2. Shape of an impacted area in a composite sheet after low-energy impact loading

lies on the opposite (i.e. inside) surface of the element wall.

The evaluation of structure polymer composites susceptibility to impact-load damage is of highly practical nature. Traditional impact tests (Charpy, Izod, et al) used in the testing of metals and polymer plastics turned out to be of limited application for the estimation of layer structure composites. Specific features of laminates are the reason why these materials need finding new methods for the evaluation of their usefulness for work in conditions which may bring about the occurrence of impact loading damage [2].

3. Experimental research

The aim of the carried out research was to determine the influence of low-energy impact load upon the mechanical properties of the composite material, i.e. laminate, used for the covering of the front part of the aircraft fuselage. The research was conducted on the basis of a programme specified by an algorithm (Fig. 3). It bases upon the assumption that delamination and cracking remain the main causes of laminate strength reduction, after impact loading [5, 9, 12, 15].

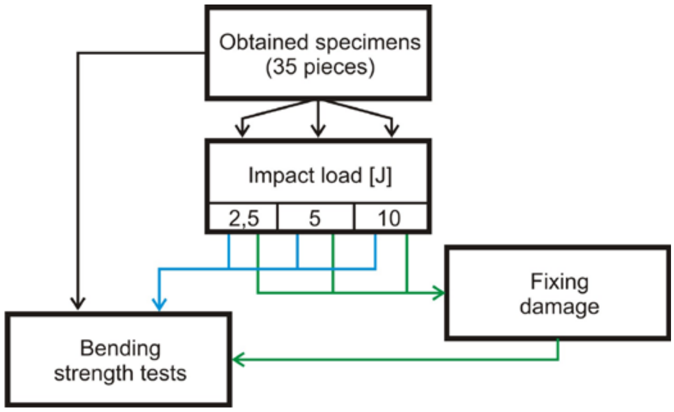


Fig. 3. Research diagram

The specimens, rectangular in shape, were obtained from an aircraft covering element – the TS-11 "Iskra", approximately 30 years in service. We paid particular attention to the lack of wear and tear as well damage emerging at a later time, during storage. The specimens were derived from places whose element curving was minimal due to the fact that the shapes had to be close to rectangular. The test material was made with layered composite on the epoxy-resin base, reinforced with three layers of glass fabric, of 300 g/m2 basis weight.

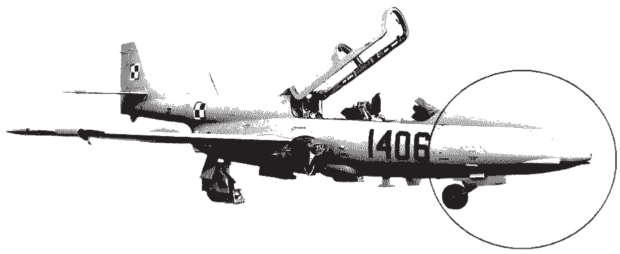


Fig. 4. The TS-11 "Iskra" aircraft; marked test element (picture by P. Idzkiewicz)

Prior to the main research we conducted a series of preliminary checks so as to select proper measurements of the specimens and the type of the impact loading. For evaluation purposes of bend strength, we adopted two most common methods, exploited for this type of examination - static expansion test [15, 16] or three-point bending test [4, 3, 13, 18].

In case of the expansion test, it was problematic to select proper specimen measurements for the research. EN ISO 527-2:1998 norm specifies the specimen width as smaller than the aperture diameter in the plates which fix the specimen on the impact loading test machine (40 mm). The specified measurements make it impossible to copy the conditions which give rise to cross-impact bend in the aircraft covering. Finally we decided to carry out a series of tests which adopted different sizes of specimens. The conducted experiments proved that the adoption of such specimens does not solve the problem as even tiny differences in specimens' thickness, (which usually occur), cause breaking of specimens in the place where they were fixed during the check, in spite of making them a bit more narrow (Fig. 5).

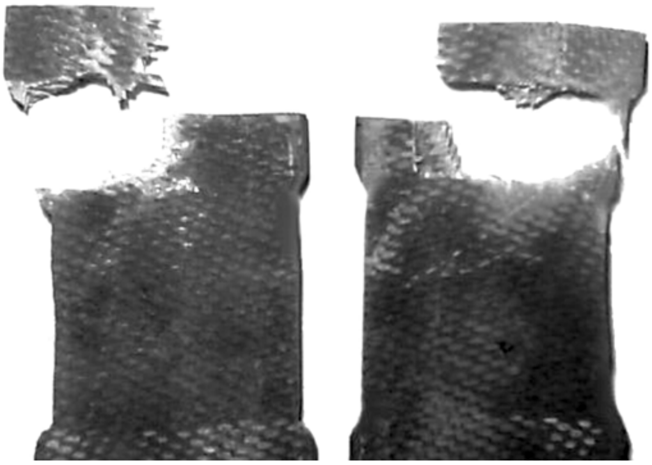


Fig. 5. Specimens for preliminary tests, broken in the fixing point

Qualifying three-point bend tests proved the suitability of the method for the assessment of the composite material bending strength.

In the first phase of the research, the specimens were impact loaded with a certain energy level, perpendicularly to the surface of the specimen. In the second phase, the specimens were checked in order to determine ultimate strength to bending.

For the sake of the research 35 specimens were used, measuring 60 x 80 mm, 4 mm thick. 5 specimens underwent tests on ultimate strength to bending so as to determine the initial strength of the tested material. Next, three series of 10 specimens, underwent cross-impact load of 2.5 J, 5 J and 10 J energy levels. Later, 5 specimens from each series were checked in order to determine the effect of the impact load upon the bending strength of the tested composite material.

In order to conduct tests, we constructed a device for cross-impact bending strength (Fig. 6). The construction of the station enables to adjust the impact energy through altering the height of the impact hammer or through changing the impact load value.

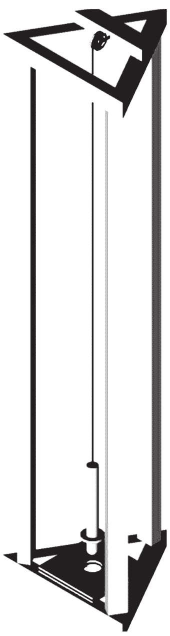


Fig. 6. Research device model

The energy impact is related to the energy conservation law. Ignoring the insignificant resistance, it was assumed that the whole potential hammer energy will be transformed into its kinetic energy at the moment of impacting the specimen. In the conducted research, we used 3 configurations of impacting energy (Table 1).

Table 1. Specification of energies and corresponding device configurations

No	Impact energy	Dart mass	Height of load
1	2.5 J	460 g	55.4 cm
2	5 J	920 g	55.4 cm
3	10 J	1460 g	69.8 cm

After performing impact load tests, we conducted a three-point bending test on a bend test machine Zwick Roell Z100. The test was done in accordance with the procedure specified in PN-EN ISO 178.

4. Damage evaluation

In case of cross-impact load of 2.5 J, the external surface of the element (specimen) does not show visible traces of the impact load; on the other hand, on the inside of the element one could notice a small delamination bulge (Fig. 7). The outer structure of the specimen did not become damaged.

After an impact load of 5 J, the external part of the specimen did

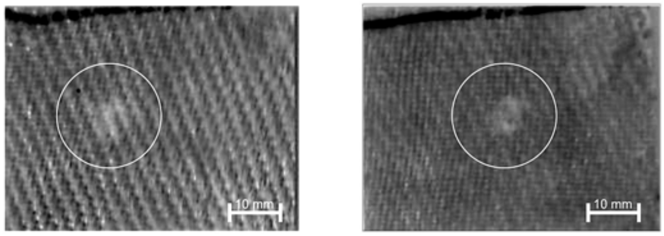


Fig. 7. Specimens after an impact load of 2.5 J (inside)

not show visible damage; however observation of the inside pointed to the emergence of a clear delamination area, with noticeable base cracking (around 8 mm in length), which had a radial shape, starting at a point of the impact load (Fig. 8). On the small area there were also damaged fibres of the reinforcement.

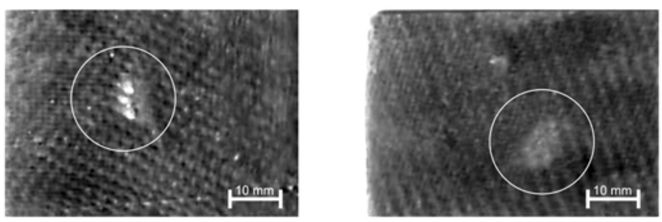


Fig. 8. Specimen after impact load of 5 J (inside)

Cross-bending impact load of 10 J caused much higher damage. On the outer side of the specimen, covered in varnish coat, there was a noticeable varnish chip. The inside surface is bulged. Both the base and the reinforcement within the radius of 10 mm from the impact loading point are damaged. Within the damaged area one can observe broken fibres of the reinforcement. Also the base materials is fragmented in certain parts. (Fig. 9).

In a series of specimens which did not undergo impact load, the character of the damage is of two different kinds (Fig. 10a).

The impact load of 2.5 J and 5 J causes a slight decrease of the bending strength. The specimens which underwent a 10 J impact load a)

b)

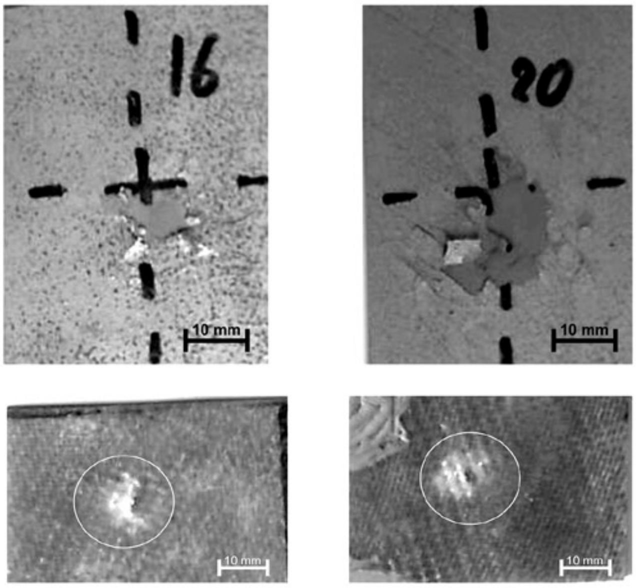


Fig. 9. Specimens after impact load of 10 J energy level: a)external side, b) inside

It is important to remember that the impact energy of 10 J equals dropping of a tool at a mass of around 1.5 kg at a height of about 70 cm. Therefore dropping a tool of a similar mass at such height can significantly lower the laminate structure.

5. Conclusions

The conducted research proved the effects of low-energy cross-impact load upon the shear strength of the composite material (laminate).

On the basis of the research one can draw the following conclusions:

1. In case of lower energy impact loading, the damage is usually difficult to locate, as the outer coat remains usually intact. It is necessary to prepare and implement non-destructive testing techniques meant to search for this type of damage for maintenance of aircraft construction elements, made with composite;
2. Destruction caused by cross-impact load causes degradation of strength properties of layer composite materials. Specimen bending strength, under the impact load of 10 J energy level, has decreased by about 37%, as referred to undamaged material.
3. The course of damage during the bending of specimens under 10 J impact load may prove that the load effects the composite

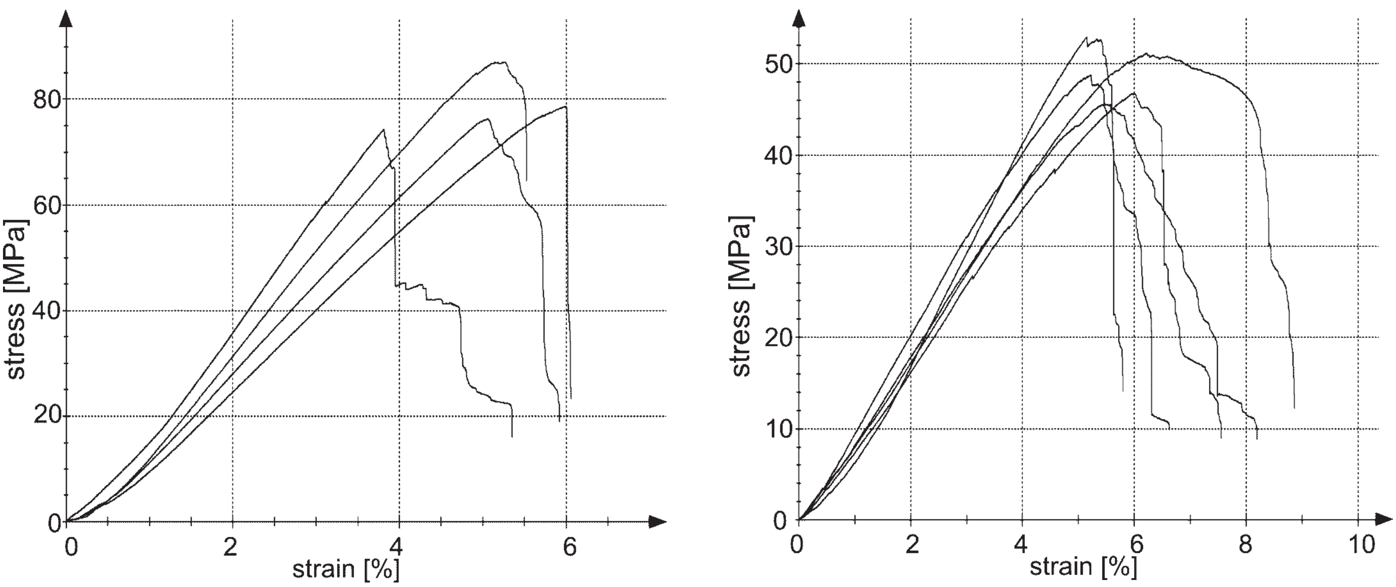


Fig. 10. Characteristics of specimens' bending strength: a) without impact load b) after 10 J energy level impact load

had the lowest bending strength. In this case, the bending strength fell by about 37% as compared to the undamaged material. Also the character of the bending curve became different (fig. 10b). On the basis of the macroscopic observation it was proved that due to a visible perforation of the laminate, the fibres which build the reinforcement material became damaged. The course of the damage, after exceeding the expansion limit Rg may prove that impact load, to a large extent, affects the material base, which leads to a gentle and gradual course of damage.

A clear difference in bending strength among the series of 5 and 10 J energy loads points to the necessity of testing the influence of cross-impact load of 5 J and 10 J, upon the bending strength properties of layered composite materials. The demand to conduct additional research is practical in its nature, since many tools used in aircraft maintenance have the mass, which after falling on the composite airplane or helicopter brings about an impact load, whose energy equals the values within the range of 5….10 J.

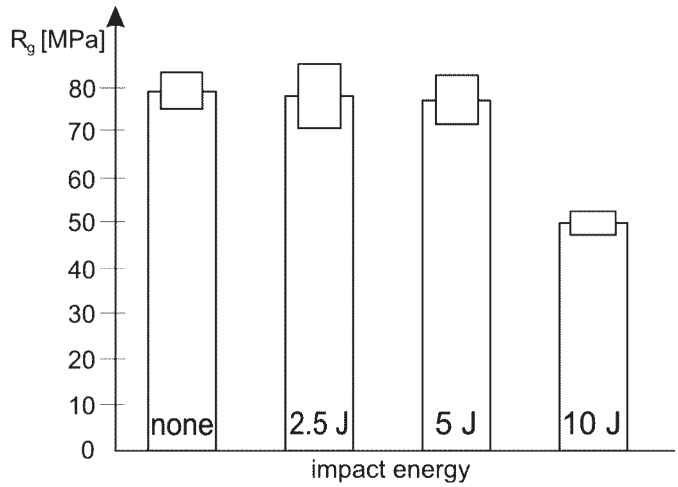


Fig. 11. Bending strength of the test specimens

base to a larger extent than in the case of undamaged material.

4. For the maintenance of aircraft made with composite material, it is vital that even a very low energy impact may endanger the flight's safety.

The carried out tests suggest that, in case of aircraft construction elements which do not bear high tensions during their service, the properties of the base material may have a tremendous importance at a time of low-energy impact damage. In the phase of designing and producing laminates used in aviation, it is important to take into account the bend-strength properties of all composite components. Properly selected material for the base will not lead to the damage of the aircraft covering structure, due to an impact load of a dropped tool, foreign bodies thrust by aircraft wheels, etc., and ultimately will not bring about an accident or an aviation catastrophe.

It seems advisable to conduct research on fatigue of composite materials under cross impact loading.

References

1. Abrate S. Impact on laminated composite materials. Applied Mechanics Reviews 1991; 44: 155-190.
2. Bełzowski A, Rechul Z, Stasieńko J. Uszkodzenia udarowe w laminacie wzmocnionym tkaniną szklaną. Kompozyty 2002; 5: 394-399.
3. Holloway L. Current development in materials technology and engineering-pultrusion. RAPRA reviews No 3 1989; 2(3).
4. Hou J P, Jerominidis G. Bending stiffness of composite plates with delamination. Elsevier. Composites: Part A 2000; 31: 121-132.
5. Kang T J, Kim C. Energy-absorption mechanisms in Kevlar multiaxial warp-knit fabric composites under impact loading. Composite Science and Technology 2000; 60: 773-784.
6. Liu D, Malvern L E. Matrix cracking in impacted glass/epoxy plates. Journal of Composite Materials 1987; 21: 594-609.
7. Mayer R W. Handbook of Pultrusion Technology. London: Chapman and Hall, 1985.
8. Mouritz A P, Gallagher J, Goodwin A A. Flexural strength and interlaminar shear strength of stitched GRP laminates following repeated impacts. Composite Science and Technology 1997; 57: 509-522.
9. Naik N K, Shrirao P, Reddy B C K. Ballistic impact behaviour of woven fabric composites: Formulation. International Journal of Impact Engineering 2006; 32(9): 1521-1552.
10. Ochelski S. Metody doświadczalne mechaniki kompozytów konstrukcyjnych. Warszawa: WNT, 2004.
11. Richardson M O W, Wisheart M J. Review of low-velocity impact properties of composite material. Elsevier Composites part A 27A 1996; 27: 1123-1131.
12. Schrauwen B, Peijs T. Influence of Matrix Ductility and Fibre Architecture on the Repeated Impact Response of Glass-Fibre-Reinforced Laminated Composites. Applied Composite Materials 2002; 9: 331-352.
13. Shim V P W, Yang L M. Characterization of the residual mechanical properties of woven fabric reinforced composites after low-velocity impact. International Journal of Mechanical Sciences 2005; 47: 647-665.
14. Shivakumar K N, Elber W, Illg W. Prediction of low-velocity impact damage in thin circular laminates. AIAA J. 1985; 23(3): 442-449.
15. da Silva Jr J E L, Paciornik S, d'Almeida J R M. Determination of the post-ballistic impact mechanical behaviour of a 45 glass-fabric composite. Polymer Testing 2004; 23: 599-604.
16. da Silva Jr J E L, Paciornik S, d'Almeida J R M. Evaluation of the effect of the ballistic damaged area on the residual impact strength and tensile stiffness of glass-fabric composite materials. Composite Structures 2004; 64: 123-127.
17. Sjoblom P O, Hartness J T, Cordell T M. On low-velocity impact testing of composite materials. Journal of Composites Materials 1988; 22: 30-52.
18. Zhang Z Y, Richardson M O W. Low velocity impact induced damage evaluation and its effect on the residual flexural properties of pultruded GRP composites. Composite Structures 2007; 81: 195-201.

Andrzej KOMOREK, Ph.D., Eng.
Paweł PRZYBYŁEK, Ph.D., Eng.
Polish Air Force Academy
ul. Dywizjonu 303 nr 12
08-521 Dęblin, Poland
E-mail: komman@op.pl; sqcdr@wp.pl

Article citation info:
JIANG G, ZHU M, WU Z. Reliability allocation using probabilistic analytical target cascading with efficient uncertainty propagation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 270–277.

Guangtai JIANG
Ming ZHU
Zhe WU

RELIABILITY ALLOCATION USING PROBABILISTIC ANALYTICAL TARGET CASCADING WITH EFFICIENT UNCERTAINTY PROPAGATION

ALOKACJA NIEZAWODNOŚCI Z WYKORZYSTANIEM PROBABILISTYCZNEJ METODY ANALITYCZNEGO KASKADOWANIA CELÓW ZAPEWNIAJĄCA WYDAJNĄ PROPAGACJĘ NIEPEWNOŚCI

Analytical target cascading (ATC) provides a systematic approach in solving reliability allocation problems for large scale system consisting of a large number of subsystems, modules and components. However, variability and uncertainty in design variables (e.g., component reliability) are usually inevitable, and when they are taken into consideration, the multi-level optimization will be very complicated. The impacts of uncertainty on system reliability are considered in this paper within the context of probabilistic ATC (PATC) formulation. The challenge is to reformulate constraints probabilistically and estimate uncertainty propagation throughout the hierarchy since outputs of subsystems at lower levels constitute inputs of subsystems at higher levels. The performance measure approach (PMA) and the performance moment integration (PMI) method are used to deal with the two objectives respectively. To accelerate the probabilistic optimization in each subsystem, a unified framework for integrating reliability analysis and moment estimation is proposed by incorporating PATC with single-loop method. It converts the probabilistic optimization problem into an equivalent deterministic optimization problem. The computational efficiency is remarkably improved as the lack of iterative process during uncertainty analysis. A nonlinear geometric programming example and a reliability allocation example are used to demonstrate the efficiency and accuracy of the proposed method.

Keywords: optimal reliability allocation, hierarchical decomposition, probabilistic analytical target cascading, uncertainty propagation.

Analityczne kaskadowanie celów (ATC) stanowi systematyczne podejście do rozwiązywania zagadnień alokacji niezawodności dotyczących systemów wielkoskalowych składających się z dużej liczby podsystemów, modułów i elementów składowych. Jednakże zmienność i niepewność zmiennych projektowych (np. niezawodności elementów składowych) są zazwyczaj nieuniknione, a gdy weźmie się je pod uwagę, optymalizacja wielopoziomowa staje się bardzo skomplikowana. W prezentowanym artykule, wpływ niepewności na niezawodność systemu rozważano w kontekście formuły probabilistycznego ATC (PATC). Wyzwanie polegało na probabilistycznym przeformułowaniu ograniczeń oraz ocenie propagacji niepewności w całej hierarchii, jako że wyjścia podsystemów na niższych poziomach stanowią wejścia podsystemów na poziomach wyższych. Cele te realizowano, odpowiednio, przy użyciu metody minimum funkcji granicznej (performance measure approach, PMA) oraz metody całkowania momentów statystycznych funkcji granicznej (performance moment integration, PMI). W celu przyspieszenia probabilistycznej optymalizacji w każdym podsystemie, zaproponowano ujednolicone ramy pozwalające na integrację analizy niezawodności z oceną momentów statystycznych poprzez połączenie PATC z metodą jednopoziomową (pojedynczej pętli, single-loop method). Zaproponowana metoda polega na przekształceniu probabilistycznego zagadnienia optymalizacyjnego na deterministyczne zagadnienie optymalizacyjne. Zwiększa to znacznie wydajność obliczeniową w związku z brakiem procesu iteratywnego podczas analizy niepewności. Wydajność i trafność proponowanej metody wykazano na podstawie przykładów dotyczących programowania nieliniowego geometrycznego oraz alokacji niezawodności.

Słowa kluczowe: *optymalna alokacja niezawodności, dekompozycja hierarchiczna, probabilistyczna metoda analitycznego kaskadowania celów, propagacja niepewności.*

1. Introduction

Optimal reliability design (reliability allocation) aims to determine the reliability of constituent subsystems and components so as to obtain targeted overall system reliability. It should be performed early in the design cycle to guide later tradeoff and improvement studies of more detailed designs. However, reliability allocation for designing complex system, such as structural, aerospace or automotive systems, is a complicated large-scale problem. Decomposition can result in improved computational efficiency because the formulation of each element typically has fewer degrees of freedom and fewer constraints than the all-in-once (AIO) formulation. Since the subsystems are coupled, their interactions need to be taken into consideration to achieve consistent designs. Zhang [20] proposed the collaborative allocation (CA) to deal with optimum allocation problem in aircraft conceptual design, which is of similar solution procedure with collaborative optimization (CO). However, in CA the auxiliary constraints are equality constraints, and the convergence has not been demonstrated yet. Recently, analytical target cascading (ATC) has been applied successfully to a variety of reliability allocation problems [4, 11, 21]. ATC is a methodology for cascading upper level design targets to lower level while the element at the lower level tries to provide responses as close to these targets as possible [8]. It has a few features which are applicable to optimum allocation problem. Firstly, upper level providing lower level with targets of variables is similar to allocation of design requirements. Secondly, the hierarchic multilevel optimization of ATC is similar to system structure composed of subsystems, components and parts. Finally, by forcing the consistency between each subsystem, ATC has proven to be convergent to the original undecomposed problem.

However, there exists uncertainty in design variables or parameters in the early development stage. For example, component reliability estimates are often uncertain, particular for new products with few failure data [2]. Thus, accurate estimates of system risk should be sought and used in system design and trade studies. In response to these new requirements, the ATC formulation has been extended to solve probabilistic design optimization problems using random variables to represent uncertainty [10], and generalized with general probabilistic characteristics by Liu [14]. In the previously published probabilistic ATC (PATC) formulations, the first few moments are usually used as targets and responses since matching two random variables is not practically doable in most cases. Even with the first few moments, however, computing the solution is very expensive due to computational difficulty in estimating propagated uncertainty. An efficient and accurate mechanism is required for propagating probabilistic information throughout the hierarchy.

$$
\begin{aligned}
&\text{Given } \boldsymbol{\mu}^{\mathrm{U}}_{\mathbf{R}_{ij}}, \boldsymbol{\sigma}^{\mathrm{U}}_{\mathbf{R}_{ij}}, \boldsymbol{\mu}^{\mathrm{U}}_{\mathbf{Y}_{ij}}, \boldsymbol{\sigma}^{\mathrm{U}}_{\mathbf{Y}_{ij}}, \boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{R}_{(i+1)k}}, \boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{R}_{(i+1)k}}, \\
&\qquad \boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{Y}_{(i+1)k}}, \boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{Y}_{(i+1)k}}, \quad k=1,\ldots,n_{ij} \\
&\text{find } \quad \boldsymbol{\mu}_{\mathbf{R}_{(i+1)k}}, \boldsymbol{\sigma}_{\mathbf{R}_{(i+1)k}}, \mathbf{X}_{ij}, \boldsymbol{\mu}_{\mathbf{Y}_{ij}}, \boldsymbol{\sigma}_{\mathbf{Y}_{ij}}, \boldsymbol{\mu}_{\mathbf{Y}_{(i+1)k}}, \\
&\qquad \boldsymbol{\sigma}_{\mathbf{Y}_{(i+1)k}}, \varepsilon_{ij}^{\boldsymbol{\mu}\mathbf{R}}, \varepsilon_{ij}^{\boldsymbol{\sigma}\mathbf{R}}, \varepsilon_{ij}^{\boldsymbol{\mu}\mathbf{Y}}, \varepsilon_{ij}^{\boldsymbol{\sigma}\mathbf{Y}}, k=1,\ldots,n_{ij} \\
&\min \quad \left\|\boldsymbol{\mu}_{\mathbf{R}_{ij}} - \boldsymbol{\mu}^{\mathrm{U}}_{\mathbf{R}_{ij}}\right\| + \left\|\boldsymbol{\sigma}_{\mathbf{R}_{ij}} - \boldsymbol{\sigma}^{\mathrm{U}}_{\mathbf{R}_{ij}}\right\| + \left\|\boldsymbol{\mu}_{\mathbf{Y}_{ij}} - \boldsymbol{\mu}^{\mathrm{U}}_{\mathbf{Y}_{ij}}\right\| \\
&\qquad + \left\|\boldsymbol{\sigma}_{\mathbf{Y}_{ij}} - \boldsymbol{\sigma}^{\mathrm{U}}_{\mathbf{Y}_{ij}}\right\| + \varepsilon_{ij}^{\boldsymbol{\mu}\mathbf{R}} + \varepsilon_{ij}^{\boldsymbol{\sigma}\mathbf{R}} + \varepsilon_{ij}^{\boldsymbol{\mu}\mathbf{Y}} + \varepsilon_{ij}^{\boldsymbol{\sigma}\mathbf{Y}} \\
&\text{s.t.} \quad \sum_{k=1}^{n_{ij}} \left\|\boldsymbol{\mu}_{\mathbf{R}_{(i+1)k}} - \boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{R}_{(i+1)k}}\right\| \le \varepsilon_{ij}^{\boldsymbol{\mu}\mathbf{R}} \\
&\qquad \sum_{k=1}^{n_{ij}} \left\|\boldsymbol{\sigma}_{\mathbf{R}_{(i+1)k}} - \boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{R}_{(i+1)k}}\right\| \le \varepsilon_{ij}^{\boldsymbol{\sigma}\mathbf{R}} \\
&\qquad \sum_{k=1}^{n_{ij}} \left\|\boldsymbol{\mu}_{\mathbf{Y}_{(i+1)k}} - \boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{Y}_{(i+1)k}}\right\| \le \varepsilon_{ij}^{\boldsymbol{\mu}\mathbf{Y}} \\
&\qquad \sum_{k=1}^{n_{ij}} \left\|\boldsymbol{\sigma}_{\mathbf{Y}_{(i+1)k}} - \boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{Y}_{(i+1)k}}\right\| \le \varepsilon_{ij}^{\boldsymbol{\sigma}\mathbf{Y}} \\
&\qquad \Pr\left(G_{ij,m}\left(\mathbf{R}_{ij}, \mathbf{X}_{ij}, \mathbf{Y}_{ij}\right) \ge 0\right) \ge P_{ij,m}, m=1,\ldots,M \\
&\text{where } \mathbf{R}_{ij} = \mathbf{f}_{ij}\left(\mathbf{R}_{(i+1)1},\ldots,\mathbf{R}_{(i+1)n_{ij}}, \mathbf{X}_{ij}, \mathbf{Y}_{ij}\right)
\end{aligned}
\tag{1}
$$

The paper proceeds as follows. First, the general PATC formulation is revisited in Section 2. Section 3 provides an introduction of existing methods for uncertainty propagation. Section 4 develops an efficient methodology integrating single loop method to deal with the issue of modeling uncertainty in multilevel hierarchies. The efficiency and accuracy of the proposed algorithm is demonstrated on two examples in Section 5. Finally, conclusions are presented in Section 6.

2. PATC formulation based on matching mean and variance

The choice of probabilistic characteristic is an important issue in decomposition based system design optimization under uncertainty because it is not practical to match two distributions exactly. For distributions with negligible higher-order moments, matching only the first two moments (mean value and variance) should be sufficient. According to the general PATC formulation provided by Liu [14], the design optimization problem for element j at level i (element O_{ij}) is shown in equation (1).

For a subsystem at certain level, its neighboring lower-level subsystems are called its children, while the neighboring upper-level subsystem is called its parent. In equation (1), $\mathbf{R}_{ij}$ and $\mathbf{Y}_{ij}$ are vectors of random responses and linking variables, respectively. $\mathbf{R}_{ij}$ are evaluated using analysis or simulation models $\mathbf{R}_{ij} = \mathbf{f}_{ij}\left(\mathbf{R}_{(i+1)1},\ldots,\mathbf{R}_{(i+1)n_{ij}}, \mathbf{X}_{ij}, \mathbf{Y}_{ij}\right.$. Targets for mean and standard deviation of $\mathbf{R}_{ij}$ and $\mathbf{Y}_{ij}$ are assigned by the parent element as $\boldsymbol{\mu}^{\mathrm{U}}_{\mathbf{R}_{ij}}$, $\boldsymbol{\sigma}^{\mathrm{U}}_{\mathbf{R}_{ij}}$ and $\boldsymbol{\mu}^{\mathrm{U}}_{\mathbf{Y}_{ij}}$, $\boldsymbol{\sigma}^{\mathrm{U}}_{\mathbf{Y}_{ij}}$. Achievable mean and standard deviation of $\mathbf{R}_{ij}$ and $\mathbf{Y}_{ij}$ are feed back to its parent element as $\boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{R}_{ij}}$, $\boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{R}_{ij}}$ and $\boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{Y}_{ij}}$, $\boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{Y}_{ij}}$. Similarly, achievable values of its children element responses and linking variables are passed to O_{ij} as $\boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{R}_{(i+1)k}}$, $\boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{R}_{(i+1)k}}$ and $\boldsymbol{\mu}^{\mathrm{L}}_{\mathbf{Y}_{(i+1)k}}$, $\boldsymbol{\sigma}^{\mathrm{L}}_{\mathbf{Y}_{(i+1)k}}$. The design consistency is formulated as the first four constrains in equation (1). The optimization problem for O_{ij} is to find the optimum values for local design variables $\mathbf{X}_{ij}$, linking variables $\mathbf{Y}_{ij}$ and the target values for responses $\boldsymbol{\mu}_{\mathbf{R}_{(i+1)k}}$, $\boldsymbol{\sigma}_{\mathbf{R}_{(i+1)k}}$ and linking variables $\boldsymbol{\mu}_{\mathbf{Y}_{(i+1)k}}$, $\boldsymbol{\sigma}_{\mathbf{Y}_{(i+1)k}}$ of its children. Generally, the optimization problem in equation (1) can be formulated as

$$
\begin{aligned}
&\min_{\mathbf{d},\mathbf{X}} f\left(\boldsymbol{\mu}_{\mathbf{R}_{ij}}(\mathbf{d},\mathbf{X},\mathbf{P}), \boldsymbol{\sigma}_{\mathbf{R}_{ij}}(\mathbf{d},\mathbf{X},\mathbf{P}), \mathbf{d}, \mathbf{X}\right) \\
&\text{s.t. } \Pr\left(G_{ij,m}(\mathbf{d},\mathbf{X},\mathbf{P}) \ge 0\right) \ge P_{ij,m} \quad m=1,\ldots,M \\
&\qquad \mathbf{R}_{ij} = \mathbf{f}_{ij}\left(\mathbf{d},\mathbf{X},\mathbf{P}\right)
\end{aligned}
\tag{2}
$$

where **d** is the vector of deterministic design variables, **X** is the vector of random design variables and **P** is the vector of random parameters. The optimization problem contains probabilistic constraints and the probability of success should be calculated. Besides, in a multilevel hierarchy, outputs of subsystems at lower levels constitute inputs of subsystems at higher levels. It is thus necessary to estimate the statistical moments of these outputs with adequate accuracy. This needs to be done for all problems at all levels of the hierarchy, and the high computational cost is a great challenge.

In previous work, the Monte Carlo simulation (MCS) is used to calculate all the probabilistic characteristics of the responses, and all probabilistic constraints are simplified into the moment-matching formulations [14]. However, computational time becomes a significant challenge. MCS may not be a practical approach for design optimization problems that require a significant number of iterations. An effective way to improve efficiency is based on the Taylor series expansions, which may introduce large approximation errors of expected values for the nonlinear responses [10]. Therefore, an appropriate uncertainty propagation method needs to be selected to achieve an appropriate balance between accuracy and efficiency.

3. Uncertainty propagation methods

One of the key components of uncertainty analysis is the quantification of uncertainties in the system output performances propagated from uncertain inputs, named as uncertainty propagation (UP) [12]. For the optimization problem in equation (2), it should be point out that the emphases of the two kinds of uncertainty calculation problems are different. One emphasizes on assessing the performance reliability. And the other focuses on evaluating the low-order moments (mean and variance) of a performance. Thus, they are discussed separately.

3.1. Reliability analysis

Reliability analysis is a tool to compute the reliability index or the probability of failure corresponding to a given failure mode or for the entire system [5]. To deal with the reliability constraints in equation (2), the reliability index is statistically defined by a cumulative distribution function $F_{G_{ij,m}}(0)$ as

$$\begin{aligned} &P\left(G_{ij,m}(\mathbf{d},\mathbf{X},\mathbf{P}) \geq 0\right) = F_{G_{ij,m}}(0) \\ &= \int_{G_{ij,m}(\mathbf{d},\mathbf{X},\mathbf{P})\geq 0} \cdots \int f_{\mathbf{X},\mathbf{P}}\left(\mathbf{X},\mathbf{P}\right) d\mathbf{X} d\mathbf{P} \geq \Phi\left(\beta_{t,m}\right) = P_{ij,m} \end{aligned} \tag{3}$$

where Φ is the cumulative distribution function for standard normal distribution and $\beta_{t,m}$ is the target reliability index. $f_{\mathbf{X},\mathbf{P}}\left(\mathbf{X},\mathbf{P}\right)$ is a joint probability density function (PDF), which needs to be integrated. There are two different methods for the reliability assessment: the reliability index approach (RIA) [16] and the performance measure approach (PMA) [18]. RIA uses the reliability index (equation (4)) to describe the probabilistic constraint in equation (3).

$$\beta_{s,m} = \left(\Phi^{-1}\left(F_{G_{ij,m}}\left(0\right)\right)\right) \geq \beta_{t,m} \tag{4}$$

where $\beta_{s,m}$ is the safety reliability index for the m[th] probabilistic constraint. In RIA, the first-order safety reliability index is obtained using first-order reliability method (FORM). It is formulated as an optimization problem, with an implicit equality constraint in a standard U space defined as the limit state function.

$$\begin{aligned} &\min \|\mathbf{U}\| \\ &\text{s.t.} \quad G_{ij,m}\left(\mathbf{U}\right) = 0 \end{aligned} \tag{5}$$

where the vector **U** represents the random variables in the standard normal space. However, RIA may yield singularity in many reliability based design optimization (RBDO) applications. Moreover, it is more efficient in evaluating the violated probabilistic constraint. If the probabilistic constraint is inactive, RIA often yields a low rate of convergence [18]. To overcome these difficulties, PMA was developed to solve the RBDO problem. In this method, the reliability constraints are stated by an R-percentile formulation as

$$\Pr\left(G_{ij,m}\left(\mathbf{d},\mathbf{X},\mathbf{P}\right) \geq G_{ij,m}^{R}\right) = R \tag{6}$$

Equation (6) indicates that the probability of $G_{ij,m}(\mathbf{d}, \mathbf{X}, \mathbf{P})$ greater than or equal to the R-percentile $G_{ij,m}^{R}$ is exactly equal to the desired reliability *R*. Instead of calculating the probability of failure directly, PMA judges whether or not a given design satisfies the probabilistic constraint with a given target reliability index R. Therefore, the original constraints that require the reliability assessment are now converted to constraints that evaluate the R-percentile. The percentile $G_{ij,m}^{R}$ can be evaluated by the inverse reliability analysis

$$\begin{aligned} &\min G_{ij,m}\left(\mathbf{U}\right) \\ &\text{s.t.} \quad \|\mathbf{U}\| = R \end{aligned} \tag{7}$$

3.2. Moment estimation

One purpose of statistical moment estimation stems from the robust design optimization, which attempts to minimize the quality loss, which is a function of the statistical mean and standard deviation [3]. The first two statistical moments of linking variables are estimated here to solve the higher-level problems and the overall multilevel design problem. Several methods are proposed to estimate the statistical moments of the output response. Monte Carlo simulation could be accurate for the moment estimation, however it requires a very large number of function evaluations. The first order Taylor series expansion has been widely used to estimate the first and second statistical moments in robust design. Nevertheless, the first order Taylor series expansion results in a large error especially when the input random variables have large variations. To overcome the shortcomings explained above, numerical integrations method have been recently proposed. The numerical integration methods rely on the principle that the first few moments of a random variable will adequately describe the complete PDF of the variable. The random variables are assumed to be statistically independent. Analytically, the statistical moments of the performance function H(**X**) can be expressed in an integration form as

$$\begin{aligned} &E\left[H\right]^{1} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} H(\mathbf{X}) f_{\mathbf{X}}\left(\mathbf{X}\right) d\mathbf{X} = \mu_{H} \\ &E\left[\left(H(\mathbf{X}) - \mu_{H}\right)\right]^{k} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left(H\left(\mathbf{X}\right) - \mu_{H}\right)^{k} f_{\mathbf{X}}\left(\mathbf{X}\right) d\mathbf{X} \end{aligned} \tag{8}$$

where $f_{\mathbf{X}}(\mathbf{X})$ is a joint PDF of the random parameters **X**. The numerical integration can be used either on the input domain or on the output domain. Since the computation of the moment could be very expensive through numerical integration on the input domain, a new formulation called performance moment integration (PMI) method is proposed for statistical moment calculation, which using numerical integration on the output domain [19]. The statistical moment calculation can be rewritten as

$$\begin{aligned} &E\left[H\right]^{1} = \int_{-\infty}^{\infty} h f_{H}\left(h\right) dh = \mu_{H} \\ &E\left[\left(H - \mu_{H}\right)\right]^{k} = \int_{-\infty}^{\infty} \left(h - \mu_{H}\right)^{k} f_{H}\left(h\right) dh \end{aligned} \tag{9}$$

where $f_H(h)$ is a probability density function of H. To approximate the statistical moments of H accurately, N-point numerical quadrature technique can be used as

$$\begin{aligned} &E\left[H\right]^{1} = \mu_{H} \cong \sum_{i=1}^{N} w_i h_i \\ &E\left[H - \mu_{H}\right]^{k} \cong \sum_{i=1}^{N} w_i (h_i - \mu_{H})^{k} \text{ for } 2 \leq k \leq 5 \end{aligned} \tag{10}$$

At minimum, the three-point integration is required to maintain a good accuracy in estimating first two statistical moments. By solving equation (10), three levels and weights on the output domain are obtained as $\{h_1, h_2, h_3\} = \left\{h_{\beta=-\sqrt{3}}, h(\mu_{\mathbf{X}}), h_{\beta=+\sqrt{3}}\right\}$ and $\{w_1, w_2, w_3\} = \left\{\frac{1}{6}, \frac{4}{6}, \frac{1}{6}\right\}$, respectively. Then, the mean and standard variation of the output response are approximated to be

$$\begin{aligned} &E\left[H\right]^{1} = \mu_{H} \cong \frac{1}{6} h_{\beta=-\sqrt{3}} + \frac{4}{6} h\left(\boldsymbol{\mu}_{\mathbf{X}}\right) + \frac{1}{6} h_{\beta=+\sqrt{3}} \\ &E\left[H - \mu_{H}\right]^{2} = \sigma_{H}^{2} = \int_{-\infty}^{\infty} \left(h - \mu_{H}\right)^{2} f_{H}\left(h\right) dh \\ &\qquad \cong \frac{1}{6}\left(h_{\beta=-\sqrt{3}} - \mu_{H}\right)^{2} + \frac{1}{6}\left(h_{\beta=+\sqrt{3}} - \mu_{H}\right)^{2} \end{aligned} \tag{11}$$

In equations (11), $h_{\beta=-\sqrt{3}}$ and $h_{\beta=+\sqrt{3}}$ can be obtained through inverse reliability analysis. The optimization problem used to approximate $h_{\beta=-\sqrt{3}}$ can be denoted as

$$\begin{aligned}&\min h(\mathbf{U})\\&\text{s.t.}\quad \|\mathbf{U}\|=\sqrt{3}\end{aligned} \tag{12}$$

The term $h_{\beta=+\sqrt{3}}$ in equations (11) can be approximated as the optimal cost obtained by maximizing $h(\mathbf{U})$ in equation (12). The term $h(\boldsymbol{\mu}_{\mathbf{X}})$ is the performance function value at the design point.

4. Single-loop method based probabilistic analytical target cascading

As mentioned above, since both the reliability analysis and the moment estimation make use of inverse reliability analysis to get percentile performance, it is very natural for the two different methodologies to be treated in a unified manner. In addition, each inverse reliability analysis is a separate optimization loop in the standard normal space. Then each subsystem optimization will be a nested, double-loop approach, which can drastically increase the computational cost. To accelerate the subsystem optimization, we employ the single loop method that has been developed for single-disciplinary systems [13]. It eliminates the need for inner reliability loops without increasing the number of design variables by using a relation representing the Karush-Kuhn-Tucker (KKT) optimality conditions instead of solving a nonlinear constrained optimization problem. The single loop method is used to efficiently evaluate percentile performances for both moment estimation and reliability assessments in PATC. The proposed strategy is named PATC-SL. For the optimization problem of equation (7), letting $R=\beta_{\mathrm{t,m}}$, the following KKT optimality condition is satisfied at the optimal point.

$$\nabla G_{ij.m}(\mathbf{U})+\lambda\nabla H(\mathbf{U})=0 \tag{13}$$

where $H(\mathbf{U})=\|\mathbf{U}\|-\beta_{t,m}$ is an equality constraint and λ is the corresponding Lagrange multiplier. According to the geometric explanation in reference [13], equation (13) states that the gradients $\nabla G_{ij,m}(\mathbf{U})$ and $\nabla H(\mathbf{U})$ are collinear and point in opposite directions. This condition yields

$$\begin{aligned}&\mathbf{U}=-\beta_{t,m}*\boldsymbol{\alpha}\\&\boldsymbol{\alpha}=\nabla G_{\mathbf{U}}(\mathbf{d},\mathbf{X},\mathbf{P})/\|\nabla G_{\mathbf{U}}(\mathbf{d},\mathbf{X},\mathbf{P})\|\end{aligned} \tag{14}$$

where $\boldsymbol{\alpha}$ is the constraint normalized gradient in U-space. Under the assumption for the PATC that the random variables are normally or can be approximated to be normally distributed [6], Equations (14) yield the following relationship between the most probable point (MPP) $\mathbf{X}_{mpp}$, $\mathbf{P}_{mpp}$ and the mean $\boldsymbol{\mu}_{\mathbf{X}}$, $\boldsymbol{\mu}_{\mathbf{P}}$.

$$\begin{aligned}&\mathbf{X}_{mpp}=\boldsymbol{\mu}_{\mathbf{X}}-\boldsymbol{\sigma}*\beta_{t,m}*\boldsymbol{\alpha}_{ij,m},\quad \mathbf{P}_{mpp}=\boldsymbol{\mu}_{\mathbf{P}}-\boldsymbol{\sigma}*\beta_{t,m}*\boldsymbol{\alpha}_{ij,m}\\&\boldsymbol{\alpha}_{ij,m}=\boldsymbol{\sigma}*\nabla G_{ij,m_{\mathbf{X},\mathbf{P}}}(\mathbf{d},\mathbf{X}_{mpp},\mathbf{P}_{mpp})/\left\|\boldsymbol{\sigma}*\nabla G_{ij,m_{\mathbf{X},\mathbf{P}}}(\mathbf{d},\mathbf{X}_{mpp},\mathbf{P}_{mpp})\right\|\end{aligned} \tag{15}$$

where $\boldsymbol{\sigma}$ is the standard deviation vector of random variables $\mathbf{X}$ and random parameters $\mathbf{P}$. Equations (15) hold for each constraint $G_{ij,m}$ of equation (2). Similarly, according to equation (12), $h_{\beta=\pm\sqrt{3}}$ are obtained through reliability analysis at $\beta=\pm\sqrt{3}$ confidence levels. The approximate MPP can be denoted by

$$\begin{aligned}&\mathbf{X}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}}=\boldsymbol{\mu}_{\mathbf{X}}-(\pm\sqrt{3}\boldsymbol{\sigma})*\boldsymbol{\alpha}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}}\\&\mathbf{P}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}}=\boldsymbol{\mu}_{\mathbf{P}}-(\pm\sqrt{3}\boldsymbol{\sigma})*\boldsymbol{\alpha}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}}\\&\boldsymbol{\alpha}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}}=\frac{\boldsymbol{\sigma}*\nabla\mathbf{f}_{ij\mathbf{X},\mathbf{P}}\left(\mathbf{d},\mathbf{X}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}},\mathbf{P}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}}\right)}{\left\|\boldsymbol{\sigma}*\nabla\mathbf{f}_{ij\mathbf{X},\mathbf{P}}\left(\mathbf{d},\mathbf{X}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}},\mathbf{P}_{\mathbf{R}_{ij},\beta=\pm\sqrt{3}}\right)\right\|}\end{aligned} \tag{16}$$

Using equations (15) and equations (16), the double-loop optimization problem in equation (2) is transformed to the following single-loop, equivalent deterministic optimization problem.

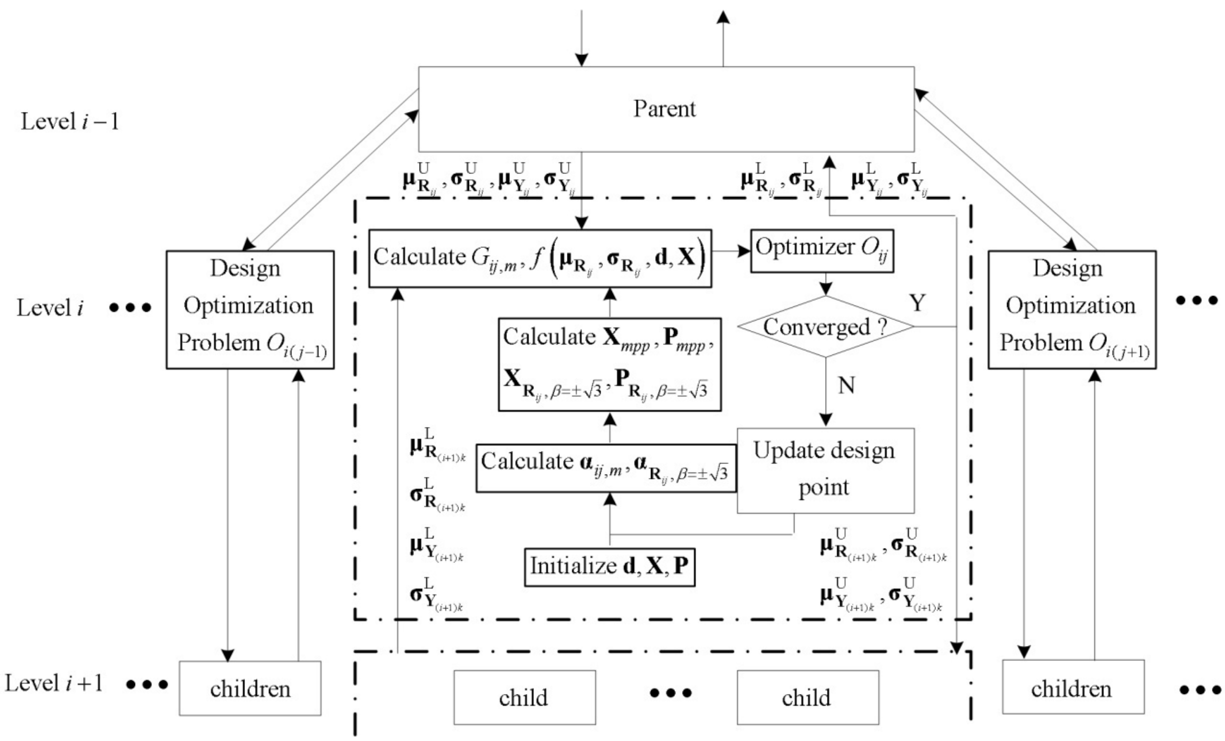


Fig. 1. Numerical process of single-loop method based PATC

$$
\begin{aligned}
&\min_{\mathbf{d},\mathbf{X}} f\left(\boldsymbol{\mu}_{\mathbf{R}_{ij}}, \boldsymbol{\sigma}_{\mathbf{R}_{ij}}, \mathbf{d}, \mathbf{X}\right)\\
&\text{s.t. } G_{ij,m}(\mathbf{d}, \mathbf{X}_{mpp}, \mathbf{P}_{mpp}) \ge 0 \quad m = 1,\ldots,M\\
&\boldsymbol{\mu}_{\mathbf{R}_{ij}} \cong \frac{1}{6}\mathbf{f}_{ij}(\mathbf{d}, \mathbf{X}_{\mathbf{R}_{ij},\beta=-\sqrt{3}}, \mathbf{P}_{\mathbf{R}_{ij},\beta=-\sqrt{3}}) + \frac{4}{6}\mathbf{f}_{ij}\left(\mathbf{d}, \boldsymbol{\mu}_{\mathbf{X}}, \boldsymbol{\mu}_{\mathbf{P}}\right)\\
&\qquad + \frac{1}{6}\mathbf{f}_{ij}(\mathbf{d}, \mathbf{X}_{\mathbf{R}_{ij},\beta=+\sqrt{3}}, \mathbf{P}_{\mathbf{R}_{ij},\beta=+\sqrt{3}})\\
&\boldsymbol{\sigma}^2_{\mathbf{R}_{ij}} \cong \frac{1}{6}\left(\mathbf{f}_{ij}(\mathbf{d}, \mathbf{X}_{\mathbf{R}_{ij},\beta=-\sqrt{3}}, \mathbf{P}_{\mathbf{R}_{ij},\beta=-\sqrt{3}}) - \boldsymbol{\mu}_{\mathbf{R}_{ij}}\right)^2\\
&\qquad + \frac{1}{6}\left(\mathbf{f}_{ij}(\mathbf{d}, \mathbf{X}_{\mathbf{R}_{ij},\beta=+\sqrt{3}}, \mathbf{P}_{\mathbf{R}_{ij},\beta=+\sqrt{3}}) - \boldsymbol{\mu}_{\mathbf{R}_{ij}}\right)^2
\end{aligned}
\tag{17}
$$

The single-loop method does not search for the MPP at each iteration. This dramatically improves the efficiency of the single-loop method without compromising the accuracy. Since $\boldsymbol{\alpha}_{ij,m}$ is a function of $\mathbf{X}_{mpp}$, equations (15) must be solved iteratively. That is, an iterative solution is obtained, where the normalized gradient from the previous iteration is used for constraint evaluation in the current iteration. The vectors $\boldsymbol{\alpha}_{ij,m}$ and $\mathbf{X}_{mpp}$ are alternately updated until the computations converge to a final probabilistic design. The same strategy is used for the calculation of $\boldsymbol{\mu}_{\mathbf{R}_{ij}}$ and $\boldsymbol{\sigma}_{\mathbf{R}_{ij}}$. Propagating uncertainty information during the PATC process should start at the bottom level of the hierarchy, where probability distribution on the input random variables and parameters are assumed as known. If such information is not available at the bottom level, start at the lowest level possible where it is available [10]. The process of PATC-SL is shown in Fig. 1.

To improve the convergence, formal methods for setting proper weights for element responses and linking variables can be found in Kim [9], Michalek [15], and Tosserams [17]. The augmented Lagrangian approach which shows stable convergence properties is used in this paper.

5. Numerical Examples

In this section, two examples are solved by the single-loop method based PATC. Comparing to other approaches, Performance of the proposed method is validated with respect to two criteria: accuracy of the solution and efficiency of the coordination process. For the accuracy comparison, the method is compared with the probabilistic all-in-one (PAIO) formulation using MCS technique (with 10000 samples), denoted as PAIO-MCS [14]. For the efficiency comparison, the current process is compared to probabilistic ATC employing linearization techniques (FORM and Taylor expansion), denoted as PATC-L [1].

5.1. Geometric programming problem

Geometric programming problem with polynomials is usually used to test the effectiveness of ATC formulations. The deterministic AIO and ATC formulations are provided by Kim [8]. Then it is formulated in a probabilistic form to demonstrate whether the PATC is capale of reaching the same optimal solution[6, 14]. The PAIO problem is formulated as

$$
\begin{aligned}
&\min E\left[f\right] = \mu_{X_1}^2 + \mu_{X_2}^2\\
&\text{with respect to}\\
&x = \left\{x_4, x_5, x_7, \mu_{X_8}, x_9, x_{10}, \mu_{X_{11}}, x_{12}, x_{13}, x_{14}\right\}^T\\
&\text{subject to} \Pr\left[g_i \le 0\right] \ge \alpha_i, \quad i = 1,\ldots,6
\end{aligned}
\tag{18}
$$

where
$$
\begin{aligned}
g_1 &= \left(X_3^{-2} + x_4^2\right)x_5^{-2} - 1\\
g_2 &= \left(x_5^2 + X_6^{-2}\right)x_7^{-2} - 1\\
g_3 &= \left(x_8^2 + x_9^2\right)X_{11}^{-2} - 1\\
g_4 &= \left(x_8^{-2} + x_{10}^2\right)X_{11}^{-2} - 1\\
g_5 &= \left(X_{11}^2 + x_{12}^{-2}\right)x_{13}^{-2} - 1\\
g_6 &= \left(X_{11}^2 + x_{12}^2\right)x_{14}^{-2} - 1\\
X_1 &= \left(X_3^2 + x_4^{-2} + x_5^2\right)^{1/2}\\
X_2 &= \left(x_5^2 + X_6^2 + x_7^2\right)^{1/2}\\
X_3 &= \left(X_8^2 + x_9^{-2} + x_{10}^{-2} + X_{11}^2\right)^{1/2}\\
X_6 &= \left(X_{11}^2 + x_{12}^2 + x_{13}^2 + x_{14}^2\right)^{1/2}
\end{aligned}
$$

Design variables X_8 and X_{11} are assumed to be independent and normally distributed with constant standard deviations $\sigma_{X_8} = \sigma_{X_{11}} = 0.1$. The required reliability level is 99.865% for all probabilistic constraints.

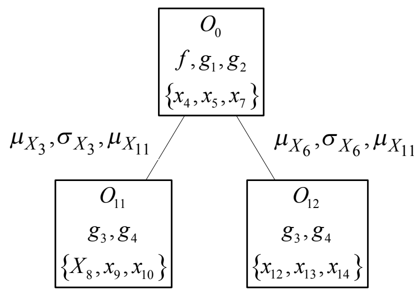


Fig. 2. Hierarchical structure of example 1

Table. 1. Optimal solutions and number of function evaluations for example 1

	PAIO-MCS	PATC-L	PATC-SL
X_4	0.76	0.754	0.76
X_5	0.86	0.855	0.86
X_7	0.91	0.905	0.906
μ_{X_8}	1.03	1.04	1.046
X_9	0.76	0.7	0.69
X_{10}	0.81	0.76	0.78
$\mu_{X_{11}}$	1.68	1.645	1.651
X_{12}	0.84	0.923	0.824
X_{13}	2.31	2.24	2.3
X_{14}	2.15	2.17	2.13
E[f]	24.67	24.9	24.7
Relative error of σ_{X_3}	0.272%	1.05%	0.397%
Relative error of σ_{X_6}	0.0437%	0.177%	0.081%
Number of function evaluations	5243×10000	40599	3305

The structure of the decomposed problem is illustrated in Fig. 2. The randomness in X_8 and X_{11} results in uncertainties in all computed response X_3, X_6, each described by its mean and standard deviation. Note that since the standard deviation of the random design variable X_{11} is assumed constant, it is not included as a linking variable. The initial point is set to the deterministic optimal point, {0.76, 0.87, 0.94, 0.97, 0.87, 0.80, 1.30, 0.84, 1.76, 1.55}.

Table 1 summarizes the results obtained from the three algorithms. The optimization algorithm used is sequential quadratic programming. The specified tolerance of consistency is 1.0×10^{-4}. σ_{X_3} and σ_{X_6} are validated via MCS with 100000 samples, and the relative errors are also displayed in Table 1. It is found that PATC-SL is more accuracy in estimating σ_{X_3} and σ_{X_6}, and achieves better solutions than PATC-L. More importantly, with no nested optimization loops, the number of function evaluations for PATC-SL is significantly smaller than those for others. Furthermore, the average number of iteration cycles of each system-level optimization using PATC-SL method is about 25, and the PATC-L method is around 70. Higher rate of convergence further improves the algorithmic efficiency. Compared to PATC-L, the results show that the PATC-SL improves the computational efficiency by more than 12 times.

5.2. Reliability optimum allocation problem

In this section, we demonstrate the methodology for reliability allocation using a two-level example. The deterministic formulation of decomposed optimization problem is presented in reference [20]. Through Fig. 3 it is apparent that the system is composed of five subsystems and each subsystem encompasses two components.

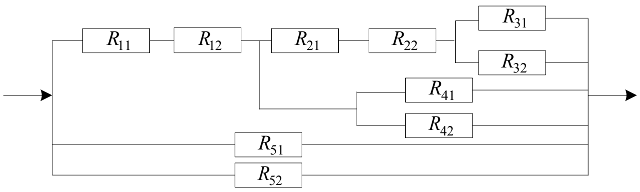


Fig. 3. The topology structure of system in the example 2

$$
\begin{aligned}
\min\ & \sum_{i=1}^{5}\sum_{j=1}^{2} C_{ij} \\
\text{s.t.}\ & R_s \ge 0.999 \\
& R_s = R_5 + R_1\left(1-R_5\right)\left(R_2R_3 + R_4 - R_2R_3R_4\right) \\
& 0.5 \le R_{ij} \le 0.98, i=1,2;\ j=1,2 \\
& 0.2 \le R_{ij} \le 0.99,\ i=3,4,5;\ j=1,2 \\
& R_i = R_{i1}R_{i2} \ge 0.5, i=1,2 \\
& 0.5 \le R_i \le 0.998, i=3,4,5 \\
& C_{i1} = R_{i1}^2/3, C_{i2} = R_{i2}^2/2, i=1,2 \\
& C_{i1} = \left[\ln(1-R_{i1})\right]^2/100, i=3,4,5 \\
& C_{i2} = \left[\ln(1-R_{i2})\right]^2/60, i=3,4,5
\end{aligned}
\tag{19}
$$

where R is the reliability requirement and C is the cost. Subscript 's', 'i' and 'ij' indicate corresponding value of main system, subsystem i and component j in subsystem i, respectively. Treating component reliability as uncertain parameters is necessary as they are usually empirically determined [7]. Then all the reliability constraints should be transformed into confidence-level formulation to ensure that system reliability requirements are met with high probability. Here, the component reliability R_{ij} follows a normal distribution. The standard deviation is assumed to be 0.005. A 95% confidence level is used for every system and subsystem reliability constraint. The corresponding PATC decomposition is shown in Fig. 4.

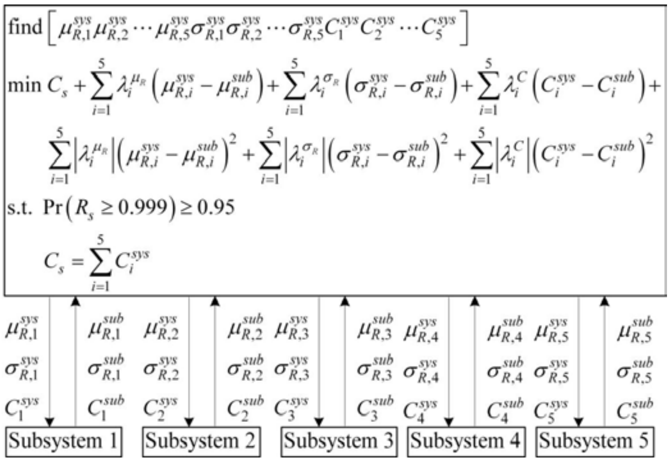


Fig. 4. The PATC-decomposed formulation of the reliability allocation problem

According to PATC, the mean and standard deviation of R_i is defined as linking variables, denoted as $\mu_{R,i}$ and $\sigma_{R,i}$ correspondingly. Auxiliary variables C_i is also introduced to calculate the total cost. Superscript '*sys*' or '*sub*' indicates the value allocated by main system or subsystem, respectively. Under the augmented Lagrangian ATC formulation [9], the consistency constraints can be incorporated into the objective function. $\lambda_i^{\mu_R}$, $\lambda_i^{\sigma_R}$ and λ_i^{C} denote the Lagrange multipliers associated with the deviations of $\mu_{R,i}$, $\sigma_{R,i}$ and C_i. The subsystem optimization can be formulated as

$$
\begin{aligned}
&\text{find } \mu_{R_{i1}}, \mu_{R_{i2}} \\
&\min\ \lambda_i^{\mu_R}\left(\mu_{R,i}^{sys}-\mu_{R,i}^{sub}\right) + \lambda_i^{\sigma_R}\left(\sigma_{R,i}^{sys}-\sigma_{R,i}^{sub}\right) + \lambda_i^{C}\left(C_i^{sys}-C_i^{sub1}\right) + \\
&\qquad \left|\lambda_i^{\mu_R}\right|\left(\mu_{R,i}^{sys}-\mu_{R,i}^{sub}\right)^2 + \left|\lambda_i^{\sigma_R}\right|\left(\sigma_{R,i}^{sys}-\sigma_{R,i}^{sub}\right)^2 + \left|\lambda_i^{C}\right|\left(C_i^{sys}-C_i^{sub1}\right)^2 \\
&\text{s.t. } \Pr\left(0.5 \le R_1^{sub1} \le 1\right) \ge 0.95 \\
&\qquad 0.5 \le \mu_{R_{i1}} \le 0.98; 0.5 \le \mu_{R_{i2}} \le 0.98, i=1,2
\end{aligned}
\tag{20}
$$

$$
\begin{aligned}
&\text{find } \mu_{R_{i1}}, \mu_{R_{i2}} \\
&\min\ \lambda_i^{\mu_R}\left(\mu_{R,i}^{sys}-\mu_{R,i}^{sub}\right) + \lambda_i^{\sigma_R}\left(\sigma_{R,i}^{sys}-\sigma_{R,i}^{sub}\right) + \lambda_i^{C}\left(C_i^{sys}-C_i^{sub1}\right) + \\
&\qquad \left|\lambda_i^{\mu_R}\right|\left(\mu_{R,i}^{sys}-\mu_{R,i}^{sub}\right)^2 + \left|\lambda_i^{\sigma_R}\right|\left(\sigma_{R,i}^{sys}-\sigma_{R,i}^{sub}\right)^2 + \left|\lambda_i^{C}\right|\left(C_i^{sys}-C_i^{sub1}\right)^2 \\
&\text{s.t. } \Pr\left(0.5 \le R_i^{sub1} \le 0.998\right) \ge 0.95 \\
&\qquad 0.2 \le \mu_{R_{i1}} \le 0.99; 0.2 \le \mu_{R_{i2}} \le 0.99, i=3,4,5
\end{aligned}
\tag{21}
$$

The means of component reliability $\mu_{R_{i1}}$ and $\mu_{R_{i2}}$ are treated as design variables. ATC is implemented first to find the deterministic optimal point, which is chosen as the initial design point of PAIO-MCS to prevent unnecessary and expensive reliability analyses for infeasible and otherwise undesirable design points.

The results are listed in Table 2 for comparison, where S_{ij} (i=1,2,…, 5,j=1,2) represents the component j in subsystem i. It shows that, the deterministic results have low confidence level once the uncertainty of the input variables is considered. With confidence-level constraints, both PAIO-MCS and PATC-SL improve the probability of meeting the reliability requirements. The accuracy of PATC-SL is excellent for this example as well. Fig. 5 shows the iteration histories for main optimization of ATC and PATC-SL. The efficiency of PATC-SL is comparable to the deterministic optimization.

Table. 2. Reliability optimum allocation results for example 2

		Subsystem(S_1)		Subsystem(S_2)		Subsystem(S_3)		Subsystem(S_4)		Subsystem(S_5)	
		S_{11}	S_{12}	S_{21}	S_{22}	S_{31}	S_{32}	S_{41}	S_{42}	S_{51}	S_{52}
ATC	R_{ij}	0.804	0.631	0.788	0.643	0.381	0.250	0.544	0.375	0.978	0.908
	C_{ij}	0.215	0.199	0.207	0.207	0.0023	0.0014	0.006	0.003	0.146	0.095
	R_i	0.507		0.51		0.535		0.715		0.998	
		Pr($Rs \geq$ 0.999)=0.249, Cs=1.083									
PAIO-MCS	$\mu_{R_{ij}}$	0.947	0.774	0.789	0.644	0.366	0.276	0.726	0.628	0.961	0.933
	C_{ij}	0.299	0.299	0.208	0.207	0.0021	0.0017	0.017	0.016	0.105	0.122
	μ_{R_i}	0.733		0.508		0.56		0.928		0.997	
		Pr($Rs \geq$ 0.999)=0.95, Cs=1.280									
PATC-SL	$\mu_{R_{ij}}$	0.929	0.788	0.741	0.686	0.341	0.274	0.881	0.385	0.950	0.948
	C_{ij}	0.288	0.310	0.183	0.235	0.0017	0.0017	0.045	0.004	0.090	0.146
	μ_{R_i}	0.732		0.516		0.521		0.927		0.997	
		Pr($Rs \geq$ 0.999)=0.952, Cs=1.304									

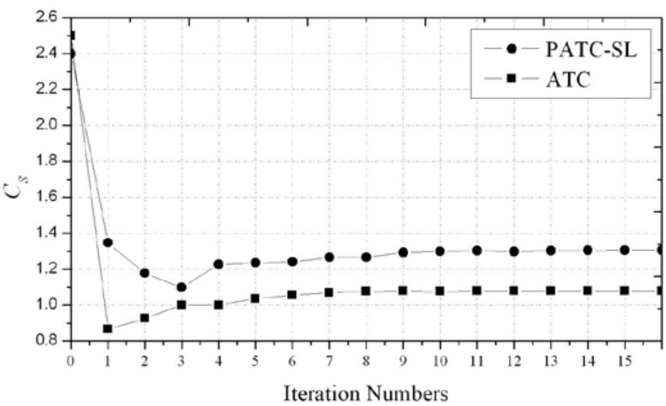


Fig. 5. Optimization history for the reliability allocation problem

6. Conclusions

The estimation uncertainty of component reliability is considered in this paper. To deal with the issue of modeling uncertainty propagation in multilevel hierarchies, PMA and PMI are investigated in the paper. A new PATC framework is developed by combining the single loop method with the uncertainty propagation techniques to solve the reliability allocation problem under uncertainty. Compared with the previous methods, the new approach requires no nested optimization loop. This makes it extremely efficient. Through the present study, it is shown that:

1) Compared to the all-in-one (AIO) method with MCS, The accuracy of the proposed PATC-SL formulation is demonstrated. The single-loop method based PATC can be useful for many nonlinear engineering systems.
2) Compared to PATC-L and ATC, the efficiency of PATC-SL is validated. Its efficiency is almost equivalent to deterministic optimization.
3) Evaluating system reliability in a probabilistic approach is meant to aid system architects make informed risk-based decisions rather than the traditional safety factor approaches. The proposed PATC-SL method is more practical for engineering application with an acceptable accuracy and better computational. Higher efficiency can be achieved by improving the convergence speed, which needs to be further studied.

References

1. Chan K Y, Kokkolaras M, Papalambros P, Skerlos S J, Mourelatos Z. Propagation of uncertainty in optimal design of multilevel systems: piston-ring/cylinder-liner case study. SAE trasactions 2004; 113(5):719–723.
2. Coit D W, Jin T, Wattanapongsakorn N. System optimization with component reliability estimation uncertainty: a multi-Criteria approach. IEEE transactions on reliability 2004; 53(3): 369–380.
3. Du X P, Chen W. Towards a better understanding of modeling feasibility robustness in engineering design. Journal of Mechanical Design 2000; 122(4):385–394.
4. Guo Z W, Bai G C. Reliability allocation using analytical target cascading method. IEEE 16th International Conference on Industrial Engineering and Engineering Management 2009.
5. Haldar A, Mahadevan S. Probability, reliability and statistical methods in engineering design. John Wiley & Sons, 2000.
6. Han J. Sequential linear programming coordination strategy for deterministic and probabilistic analytical target cascading. Ph.D. dissertation, Michigan: University of Michigan, 2008.
7. Hassan R, Crossley W. Spacecraft reliability-based design optimization under uncertainty including discrete variables. Journal of Spacecraft and Rockets 2008; 45(2): 394–405.
8. Kim H M. Target cascading in optimal system design. Ph.D. dissertation, Michigan: University of Michigan, 2001.

9. Kim H M, Chen W, Wiecek M, 2005. Lagrangian coordination for enhancing the convergence of analytical target cascading. AIAA Journal 2006; 44(10): 2197–2207.
10. Kokkolaras M, Mourelatos Z P, Papalambros P Y. Design optimization of hierarchically decomposed multilevel systems under uncertainty. Proceedings of 2004 ASME Design Engineering Technical Conferences: 30th Design Automation Conference 2004.
11. Kokkolaras M. Reliability allocation in probabilistic design optimization of decomposed systems using analytical target cascading. 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference 2006.
12. Lee S H, Chen W. A comparative study of uncertainty propagation methods for black-box type functions. Proceedings of the ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference 2007.
13. Liang J, Mourelatos Z P, Tu J. A single-loop method for reliability-based design optimization. Proceedings of ASME Design Engineering Technical Conferences 2004.
14. Liu H, Chen W, Kokkolaras M, Papalambros P, Kim H M. Probabilistic analytical target cascading: a moment matching formulation for multilevel optimization under uncertainty. Journal of Mechanical Design 2006;128(4):991–1000.
15. Michalek J J, Papalambros P Y. An efficient weighting update method to achieve acceptable inconsistency deviation in analytical target cascading. Journal of Mechanical Design 2005; 127(2): 206–214.
16. Reddy M V, Granhdi R V, Hopkins D A, Reliability based structural optimization: a simplified safety index approach. Computers and Structures 1994; 53(6):1407–1418.
17. Tosserams S, Kokkolaras M, Etman L F P, Rooda J E. Extension of analytical target cascading using augmented lagrangian coordination for multidisciplinary design. 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference 2008.
18. Tu J, Choi K K, Park Y H, A new study on reliability-based design optimization. Journal of Mechanical Design 1999; 121(4):557–564.
19. Youn B D, Choi K K, Yi K Y. Performance moment integration (PMI) method for quality assessment in reliability-based robust design optimization. Mechanics Based Design of Structures and Machines 2005; 33(2): 185-213.
20. Zhang K S, Li W J, Wei H Y. A new method for optimum allocation of design requirements in aircraft conceptual design. Chinese Journal of Aeronautics 2006; 19(3): 203–211.
21. Zhang X L, Huang H Z, Liu Y. A hierarchical decomposition approach for large system reliability allocation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2009; 3: 32–37.

Guangtai JIANG, Ph.D. Candidate
Ming ZHU, Instructor
Prof. Zhe WU, Ph.D.
School of Aeronautics and Engineering
Beijing University of Aeronautics and Astronautics
Beijing, 100191, P. R. China
E-mail: jianggt986@sina.com, zm_buaa@163.com, wuzhe@buaa.edu.cn

Article citation info:
KUMAR J, KADYAN M S, MALIK S C. Cost analysis of a two-unit cold standby system subject to degradation, inspection and priority. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 278–283.

Jitender KUMAR
Mukender Singh KADYAN
Suresh Chander MALIK

COST ANALYSIS OF A TWO-UNIT COLD STANDBY SYSTEM SUBJECT TO DEGRADATION, INSPECTION AND PRIORITY

ANALIZA KOSZTÓW DWU-ELEMENTOWEGO SYSTEMU Z REZERWĄ ZIMNĄ Z UWZGLĘDNIENIEM DEGRADACJI, KONTROLI STANU SYSTEMU ORAZ PRIORYTETOWOŚCI ZADAŃ

The present paper deals with a reliability model incorporating the idea of degradation, inspection and priority. The units may fail completely directly from normal mode. There is a single server who visits the system immediately when required. The original unit undergoes for repair upon failure while only replacement of the duplicate unit is made by similar new one. The original unit does not work as new after repair and so called degraded unit. The system is considered in up-state if any one of new/duplicate/degraded unit is operative. The server inspects the degraded unit at its failure to see the feasibility of repair. If repair of the degraded unit is not feasible, it is replaced by new one similar to the original unit in negligible time. The priority for operation to the new unit is given over the duplicate unit. The distribution of failure time follow negative exponential where as the distributions of inspection, repair and replacement times are assumed as arbitrary. The system is observed at suitable regenerative epochs by using regenerative point technique to evaluate mean time to system failure (MTSF), steady-state availability, busy period and expected number of visits by the server. A particular case is considered to see graphically the trend of mean time to system failure (MTSF), availability and profit with respect to different parameters.

Keywords: *degradation, inspection, priority, profit analysis.*

Niniejsza praca dotyczy modelu niezawodności uwzględniającego zagadnienia degradacji, kontroli stanu oraz priorytetowości zadań. Elementy mogą ulegać całkowitemu uszkodzeniu bezpośrednio z trybu normalnego. Istnieje jeden konserwator, który odwiedza system, gdy tylko zachodzi taka potrzeba. W przypadku uszkodzenia, element oryginalny podlega naprawie, podczas gdy element zapasowy (duplikat) podlega jedynie wymianie na nowy, podobny. Po naprawie, element oryginalny nie działa już jako element nowy lecz jako element zdegradowany. System uważa się za zdatny jeżeli pracuje którykolwiek z trzech typów elementów: nowy/rezerwowy/zdegradowany. W przypadku uszkodzenia elementu zdegradowanego, konserwator przeprowadza kontrolę stanu elementu, aby stwierdzić możliwość realizacji naprawy. Jeżeli naprawa elementu zdegradowanego jest niemożliwa, zostaje on wymieniony, w czasie pomijalnym, na element nowy, podobny do elementu oryginalnego. Nowy element uzyskuje priorytet pracy w stosunku do elementu rezerwowego. Rozkład czasu uszkodzenia jest rozkładem wykładniczym ujemnym, a rozkłady czasów kontroli stanu, naprawy i wymiany przyjmuje się jako rozkłady dowolne. System obserwuje się w odpowiednich okresach odnowy wykorzystując technikę odnowy RPT (regenerative point technique) w celu ocenienia średniego czasu do uszkodzenia systemu (MTSF), gotowości stacjonarnej, okresu zajętości oraz oczekiwanej liczby wizyt konserwatora. Przebiegi MTSF, gotowości i zysków w funkcji różnych parametrów przedstawiono w formie graficznej na podstawie studium przypadku.

Słowa kluczowe: *degradacja, kontrola stanu, priorytetowość, analiza zysków.*

Introduction

Two-unit systems have attracted the attention of many scholars and reliability engineers for their applicability in their respective fields. A bibliography of the work on the two-unit system is given by Osaki and Nakagawa [8], Kumar and Agarwal [4]. Sridharan and Mohanavadivu [9] studied the stochastic behavior of a two-unit cold standby redundant system. But no attention was paid to reliability evaluation of cold standby system due to degradation after failure. Mokaddis et al. [7] have proposed reliability model for two-unit warm standby systems subject to degradation.

Also, sometimes repair of the degraded unit is not feasible due to its excessive use and increased cost of maintenance. In such cases, the failed degraded unit may be replaced by new one in order to avoid the unnecessary expenses of repair and this can be revealed by inspection. Malik et al. [6], Malik and Chand [5] and Kadyan et al. [2] carried out the cost-benefit analysis of systems subject to degradation with inspection for feasibility of repair. Besides, it becomes necessary to give priority in operation to new one over the duplicate unit in order to increase the reliability, availability and profit of the system. The system of non-identical units with priority for operation and repair has been discussed by Chander [1].

Keeping above facts in view, the present paper deals with a reliability model incorporating the idea of degradation, inspection and priority. The units may fail completely directly from normal mode. There is a single server who visits the system immediately when required. The original unit undergoes for repair upon failure while only replacement of the duplicate unit is made by similar new one. The original unit does not work as new after repair and so called degraded unit. The system is considered in up-state if any one of new/duplicate/degraded unit is operative. The server inspects the degraded unit at its failure to see the feasibility of repair. If repair of the degraded unit is not feasible, it is replaced by new one similar to the original unit in negligible time. The priority for operation to the new unit is given over the duplicate unit. The distribution of failure time follow negative exponential where as the distributions of inspection, repair and replacement times are assumed as arbitrary. The system is observed at suitable regenerative epochs by using regenerative point technique to evaluate mean time to system failure (MTSF), steady-state availability, busy period and expected number of visits by the server. A particular case is considered to see graphically the trend of mean time to system failure (MTSF), availability and profit with respect to different parameters.

The systems of electric transformer can be cited as a good example of the present system model.

Notation

E	:	Set of regenerative states
No	:	The unit is new and operative
NDo	:	The unit is duplicate and operative
Do	:	The unit is degraded and operative
NCs / DCs/ NDCs	:	The new/degraded/duplicate unit in cold standby
p/q	:	Probability that repair of degraded unit is feasible/not feasible
$\lambda/\lambda_1/\lambda_2$	:	Constant failure rate of new /duplicate/degraded unit
$g(t)/G(t)$, $g_1(t)/G_1(t)$	:	pdf/cdf of repair time for new/degraded unit
w(t)/W(t)	:	pdf/cdf of replacement time of the duplicate unit
h(t)/H(t)	:	pdf/cdf of inspection time of the degraded unit
$NF_{ur}/NF_{UR}/NF_{wr}$	:	New unit is failed and under repair/under continuous repair from previous state/waiting for repair.
NDF_{ure}/NDF_{URe} /NDF_{wre}/NDF_{WRe}	:	Duplicate unit is failed and under replacement/under continuous replacement from previous state/waiting for replacement/ continuously waiting for replacement from previous state.
DF_{ur}/DF_{UR}	:	Degraded unit is failed and under repair/under repair continuously from previous state.
$DF_{ui}/DF_{wi}/DF_{UI}$	:	Degraded unit is failed and under inspection /waiting for inspection/under inspection continuously from the previous state.
$q_{ij}(t),Q_{ij}(t)$	:	pdf and cdf of first passage time from regenerative state i to a regenerative state j or to a failed state j without visiting any other regenerative state in (0,t].
$q_{ij.kr}(t),Q_{ij.kr}(t)$	:	pdf and cdf of first passage time from regenerative state i to a regenerative state j or to a failed state j visiting state k,r once in (0,t].
$M_i(t)$	:	P[system up initially in state S_i ε E is up at time t without visiting any other regenerative sate]
$W_i(t)$	:	P[server is busy in the state S_i up to time t without making any transition to any other regenerative state or returning to the same via one or more non-regenerative states]
m_{ij}	:	Contribution to mean sojourn time in state $S_i \in E$ and non regenerative state if occurs before transition to $S_j \in E$.
®/©	:	Symbols for Stieltjes convolution/Laplace convolution
~|*	:	Symbols for Laplace Stieltjes Transform (LST)/Laplace Transform (LT)
'(desh)	:	Symbol for derivative of the function

The following are the possible transition states of the system model:

$S_0 = (No, NDCs)$, $S_1 = (NDo, NF_{ur})$, $S_2 = (NDF_{wre}, NF_{UR})$,
$S_3 = (NDo, DCs)$, $S_4 = (Do, NDF_{ure})$, $S_5 = (DF_{wi}, NDF_{URe})$,
$S_6 = (Do, NDCs)$, $S_7 = (NDo, DF_{ui})$ $S_8 = (NDo, DF_{ur})$, (1)
$S_9 = (NDF_{wre}, DF_{UI})$, $S_{10} = (NDF_{wre}, DF_{UR})$, $S_{11} = (No, NDF_{ure})$,
$S_{12} = (NDF_{WRe}, DF_{ur})$, $S_{13} = (NF_{wr}, NDF_{URe})$,

The states S_0, S_1, S_3, S_4, S_6 S_7, S_8 and S_{11} are regenerative states while S_2, S_5, S_9, S_{10}, S_{12}, and S_{13} are non-regenerative states. Thus $E = \{S_0, S_1, S_3, S_4, S_6\ S_7, S_8, S_{11}\}$.The possible transition between states along with transition rates for the model is shown in figure 1.

Transition Probabilities and Mean Sojourn Times

Simple probabilistic considerations yield the following expressions for the non-zero elements $p_{ij} = Q_{ij}(\infty) = \int q_{ij}(t)\,dt$ as:

$$\begin{array}{lll} p_{01} = p_{34} = p_{67}, & p_{12} = 1-g^*(\lambda_1) = p_{14.2}, & p_{13} = g^*(\lambda_1), \\ p_{46} = w^*(\lambda_2), & p_{47.5} = 1- w^*(\lambda_2) = p_{45}, & p_{7,0} = q\, h^*(\lambda_1), \\ p_{7,8} = p\, h^*(\lambda_1), & p_{7,9} = 1- h^*(\lambda_1), & p_{7,11.9} = [1- h^*(\lambda_1)]q, \\ p_{7,4.9,12} = p[1- h^*(\lambda_1)], & p_{8,3} = g_1^*(\lambda_1), & p_{8,10} = 1- g_1^*(\lambda_1) = p_{8,4.10}, \\ p_{11,0} = w^*(\lambda), & p_{11,13} = 1- w^*(\lambda) = p_{11,1.13} & \end{array} \tag{2}$$

For these transition probabilities, it can be verified that

$$p_{01}=p_{34}=p_{67}=p_{12}+p_{13}=p_{14.2}+p_{13}=p_{45}+p_{46}=p_{46}+p_{47.5}=p_{7,0}+p_{7,8}+p_{7,9}=p_{7,0}+p_{7,8}+p_{7,11.9}+p_{7,4.9,12}=p_{83}+p_{8,10}=p_{83}+p_{8,4.10}=p_{11,0}+p_{11,13}=p_{11,0}+p_{11,1.13}=1 \tag{3}$$

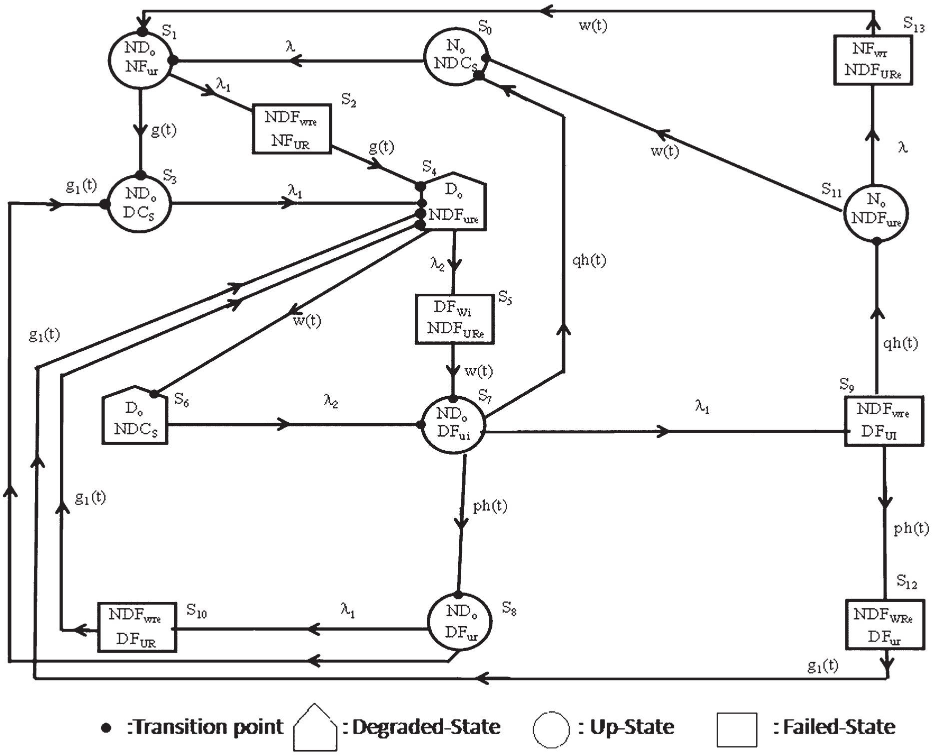


Fig. 1. State transition diagram

The mean sojourn times μ_i in state S_i are given by

$$\begin{array}{llll} \mu_0 = \dfrac{1}{\lambda}, & \mu_1 = \dfrac{1}{\lambda_1}[1-g^*(\lambda)], & \mu_3 = \dfrac{1}{\lambda_1}, & \mu_4 = \dfrac{1}{\lambda_2}[1- w^*(\lambda_2)], \\ \mu_6 = \dfrac{1}{\lambda_2}, & \mu_7 = \dfrac{1}{\lambda_2}[1-h^*(\lambda_1)], & \mu_8 = \dfrac{1}{\lambda_1}[1-g_1^*(\lambda_1)], & \mu_{11} = \dfrac{1}{\lambda}[1-w^*(\lambda)] \end{array} \tag{4}$$

The unconditional mean time taken by the system to transit from any state S_i when time is counted from epoch at entrance into state S_j is stated as:

$$m_{ij} = \int t\, dQ_{ij}(t) = -q_{ij}^{*\prime}(0) \text{ and } \mu_i = E(T) = \int_0^\infty P(T > t)dt = \sum_j m_{ij} \tag{5}$$

where T denotes the time to system failure.

Relationship Between Unconditional Mean and Mean Sojourn Times

$$\begin{array}{lll} m_{01}=\mu_0, & m_{12}+m_{13}=\mu_1, & m_{13}+m_{14.2}=\mu^1_1 \text{ (say)}, \\ m_{34}=\mu_3, & m_{45}+m_{46}=\mu_4, & m_{46}+m_{47.5}=\mu^1_4 \text{ (say)}, \\ m_{67}=\mu_6, & m_{7,8}+m_{7,10}+m_{7,9}=\mu_7, & m_{7,8}+m_{7,0}+m_{7,11.9}+m_{7,4.9,12}=\mu^1_7 \text{(say)}, \\ m_{83}+m_{8,10}=\mu_8, & m_{83} + m_{8,4.10} = \mu^1_8 \text{ (say)}, & m_{11,13}+m_{11,0}=\mu_{11}, \\ m_{11,0}+m_{11,1.13}=\mu^1_{11} \text{(say)} & & \end{array} \tag{6}$$

Mean Time to System Failure

Let $\phi_i(t)$ be the cdf of the first passage time from regenerative state i to a failed state. Regarding the failed state as absorbing state, we have the following recursive relations for $\phi_i(t)$:

$$\varphi_i(t) = \sum_j Q_{i,j}(t) ® \varphi_j(t) + \sum_k Q_{i,k}(t) \tag{7}$$

where j is an operative regenerative state to which the given regenerative state i can transit and k is a failed state to which the state i can transit directly.

Taking L.S.T. of relations (7) and solving for $\tilde{\varphi}_0(s)$.

Using this, we have

$$R^*(s) = (1-\tilde{\varphi}_0(s))/s \tag{8}$$

The reliability R(t) can be obtained by taking Laplace inverse transform of (8).

The mean time to system failure can be given by

$$\text{MTSF}(T_1) = \lim_{s\to 0} R^*(s) = \frac{N_{11}}{D_{11}} \tag{9}$$

where

$N_{11}=(1-p_{78}p_{83}p_{46})(\mu_0+\mu_1)+\mu_3[p_{13}+p_{46}(p_{83}+p_{78}p_{8,10})]+p_{13}(\mu_4+p_{46}(\mu_6+\mu_7+\mu_8 p_{78}))$

and

$D_{11}=1-p_{46}(p_{78}p_{83}+p_{13}p_{70})$

Availability Analysis

Let $A_i(t)$ be the probability that the system is in up state at instant t given that the system entered regenerative state i at $t=0$. The recursive relations for $A_i(t)$ are given by:

$$A_i(t) = M_i(t) + \sum_j q_{i,j}^{(n)}(t) © A_j(t) \tag{10}$$

where j is any successive regenerative state to which the regenerative state i can transit through $n \geq 1$ (natural number) transitions.

$$\begin{array}{llll} M_0(t)=e^{-\lambda t} & M_1(t)=e^{-\lambda_1 t}\,\bar{G}(t), & M_3(t)=e^{-\lambda_1 t}, & M_4(t)=e^{-\lambda_2 t}\,\overline{W}(t), \\ M_6(t)=e^{-\lambda_2 t}, & M_7(t)=e^{-\lambda_1 t}\,\overline{H}(t), & M_8(t)=e^{-\lambda_1 t}\,\bar{G}_1(t), & M_{11}(t)=e^{-\lambda t}\,\overline{W}(t), \end{array} \tag{11}$$

We have,

Taking LT of relations (10) and solving for $A_0^*(s)$.

The steady-state availability of the system can be given by

$$A_0(\infty) = \lim_{s\to 0} s A_0^*(s) = \frac{N_{12}}{D_{12}} \tag{12}$$

where

$N_{12}=[p_{11.9}+p_{70}]\,(\mu_0+\mu_1+p_{13}\mu_3)+\mu_4+p_{46}\mu_6+\mu_7+p_{78}(\mu_3 p_{83}+\mu_8)+p_{7,11.9}\mu_{11}$

$D_{12}=[p_{70}(\mu_0+\mu^1_1+\mu_3)+\mu^1_4+\mu_6 p_{46}+\mu^1_7+p_{78}(p_{83}\mu_3+\mu^1_8)+p_{7,11.9}(p_{11,0}\mu_0+\mu^1_1+\mu_3+\mu^1_{11})$

Busy Period Analysis for Server

Let $B_i(t)$ be the probability that the server is busy at an instant t given that the system entered regenerative state i at $t = 0$. The following are the recursive relations for $B_i(t)$

$$B_i(t) = W_i(t) + \sum_j q_{i,j}^{(n)}(t) © B_j(t) \tag{13}$$

where j is a subsequent regenerative state to which state i transits through $n \geq 1$ (natural number) transitions.

We have,

$$\begin{aligned} &W_1(t)=[e^{-\lambda_1 t}+(\lambda_1 e^{-\lambda_1 t}©1)]\,\bar{G}(t),\ W_4(t)=[e^{-\lambda_2 t}+(\lambda_2 e^{-\lambda_2 t}©1)]\,\overline{W}(t) \\ &W_7(t)=[e^{-\lambda_1 t}+(\lambda_1 e^{-\lambda_1 t}©1)]\,\overline{H}(t)+(\lambda_1 e^{-\lambda_1 t}©ph(t)©1)\,\bar{G}_1(t), \\ &W_8(t)=[e^{-\lambda_1 t}+(\lambda_1 e^{-\lambda_1 t}©1)]\,\bar{G}_1(t),\ W_{11}(t)=[e^{-\lambda t}+[(\lambda e^{-\lambda t}©1)]\,\overline{W}(t) \end{aligned} \tag{14}$$

Taking LT of relations (13) and solving for $B_0^*(s)$ and using this, we can obtain the fraction of time for which the repairman is busy in steady state

$$B_0 = \lim_{s\to 0} s B_0^*(s) = \frac{N_{13}}{D_{12}} \tag{15}$$

$N_{13}=[p_{11.9}+p_{70}]\,W_1^*(0)+W_4^*(0)+W_7^*(0)+p_{78}W_8^*(0)+p_{7,11.9}W_{11}^*(0)$

and D_{12} is already mentioned**.**

Expected Number of Visits

Let $N_i(t)$ be the expected number of visits by the server in (0,t] given that the system entered the regenerative state i at $t=0$. We have the following recursive relations for $N_i(t)$:

$$N_i(t) = \sum_j Q_{i,j}(t) ® \left[d_j + N_j(t)\right] \tag{16}$$

where j is any regenerative state to which the given regenerative state i transits and $d_i=1$, if j is the regenerative state where the server does job afresh otherwise $d_i = 0$.

Taking LST of relations (16) and solving for $\tilde{N}_0(s)$.

The expected number of visits per unit time are given by,

$$N_0 = \underset{s\to 0}{Lt}\, s\,\tilde{N}_0(s) = \frac{N_{14}}{D_{12}} \tag{17}$$

where

$N_{14}=[p_{11.9}+p_{70}](1+p_{13})+p_{46}+p_{78}p_{83}$

and D12 is already specified.

Profit Analysis

Profit incurred to the system model in steady state is given by

$$P_1=K_1A_0-K_2B_0-K_3N_0$$

Where: K_1 = Revenue per unit up time of the system

K_2 = Cost per unit time for which server is busy

K_3 = Cost per visit by the server

Particular Case

Let us take $g(t)=\theta e^{-\theta t}$, $g_1(t)=\theta_1 e^{-\theta_1 t}$, $h(t)=\alpha e^{-\alpha t}$ and $w(t)=\beta e^{-\beta t}$

By using the non-zero elements p_{ij}, we get the following results:

$MTSF(T_1)=N_{11}/D_{11}$, Availability$(A_0)=N_{12}/D_{12}$
Busy Period$(B_0)=N_{13}/D_{12}$, Expected no. of visits$(N_0)=N_{14}/D_{12}$
where

$$D_{11}=[(\theta+\lambda_1)(\alpha+\lambda_1)(\lambda_1+\theta_1)(\beta+\lambda_2)-\alpha[p\theta_1(\theta+\lambda_1)+q\theta(\lambda_1+\theta_1)]]/(\theta+\lambda_1)(\alpha+\lambda_1)(\lambda_1+\theta_1)(\beta+\lambda_2)$$

$$N_{11}=[\lambda_2\lambda_1[(\alpha+\lambda_1)(\lambda_1+\theta_1)(\beta+\lambda_2)-\beta\alpha p\theta_1]][(\lambda+\theta+\lambda_1)+\lambda_2\lambda[\theta(\alpha+\lambda_1)(\lambda_1+\theta_1)(\beta+\lambda_2)$$

$$+\beta(\theta+\lambda_1)(\theta_1(\alpha+\lambda_1)+p\alpha\lambda_1)]+\lambda\lambda_1[\lambda_2(\alpha+\lambda_1)(\lambda_1+\theta)+\beta\{(\lambda_1+\theta_1)(\alpha+2\lambda_1)+p\alpha\lambda_2\}]] \quad /[\lambda\lambda_1\lambda_2(\theta+\lambda_1)(\alpha+\lambda_1)(\lambda_1+\theta_1)(\beta+\lambda_2)]$$

$$D_{12}=[q\lambda_2\theta_1\beta\alpha(\beta+\lambda_2)(\lambda_1+\theta_1)(\beta+\lambda)(\lambda_1(\theta+\lambda)+\theta\lambda)+\theta\lambda\theta_1\lambda_1(\alpha+\lambda_1)(\lambda_1+\theta_1)(\beta+\lambda)$$

$$(\lambda_2(1+\beta A)(\beta+\lambda_2)+\beta^2)+p\lambda_2\beta\theta\lambda\alpha(\beta+\lambda)(\beta+\lambda_2)(\theta_1^2+\lambda_1(\theta_1+\lambda_1))+q\theta_1\lambda_2\lambda_1(\beta+\lambda_2)$$

$$(\lambda_1+\theta_1)(\beta^2\lambda_1\theta+(\theta+\lambda_1)\lambda\beta(\beta+\lambda)+\lambda\theta\lambda_1(\beta+\lambda))]/[\theta_1\lambda_1\theta\lambda\beta\lambda_2(\beta+\lambda_2)(\lambda_1+\theta_1)(\beta+\lambda)(\alpha+\lambda_1)]$$

$$N_{12}=[(\beta+\lambda)\{q\lambda_2(\alpha+\lambda_1)(\lambda+\lambda_1)+\lambda\lambda_1(\alpha+\lambda_1+\lambda_2)\}+\lambda_2\lambda(p\alpha(\beta+\lambda)+\lambda_1^2q)]$$

$$/[\lambda\lambda_1\lambda_2(\alpha+\lambda_1)(\beta+\lambda)$$

$$N_{13}=[(\beta q+\theta+B\beta\theta)(\alpha+\lambda_1)\theta_1+p\theta\beta\alpha+q\lambda_1\theta_1\theta]/[\theta\theta_1\beta(\alpha+\lambda_1)]$$

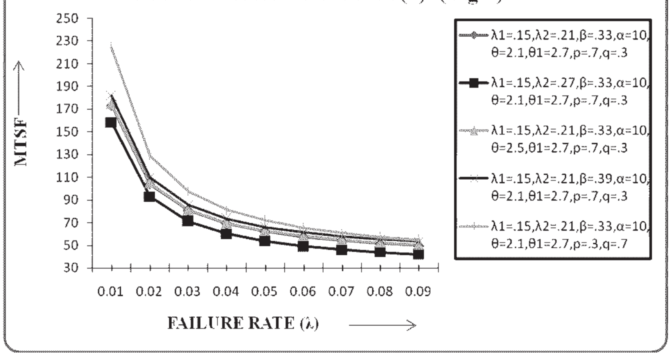
$$N_{14}=[(\alpha+\lambda_1)(\theta_1+\lambda_1)[q(2\theta+\lambda_1)(\beta+\lambda_2)+\beta(\theta+\lambda_1)]+p\alpha\theta_1(\theta+\lambda_1)(\beta+\lambda_2)]/$$

$$(\beta+\lambda_2)(\lambda_1+\theta_1)(\alpha+\lambda_1)(\theta+\lambda_1)$$

$$A=[q(\alpha+\lambda_1)^2+\alpha^2p]\theta_1\alpha+q\alpha[(\alpha+\theta_1)(\alpha+\lambda_1)^2-\alpha^2(\alpha+\theta_1+\lambda_1)]/[(\theta_1\alpha^2(\alpha+\lambda_1)^2]$$

$$B=[\theta_1(\alpha+\theta_1)(\theta_1+\lambda_1)+\alpha^2p\lambda_1]/[(\theta_1\alpha(\alpha+\theta_1)(\theta_1+\lambda_1)].$$

MTSF vs FAILURE RATE(λ) (Fig.2)

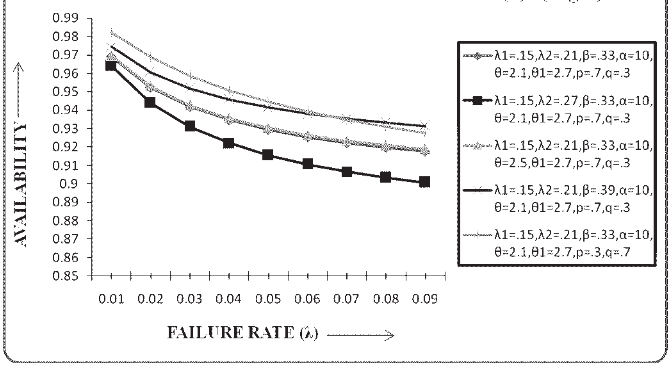


Conclusion

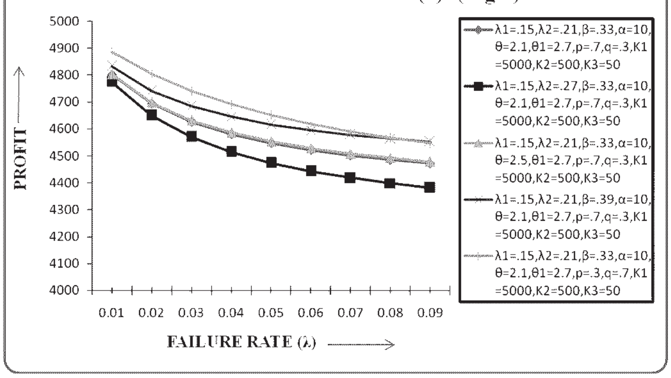
The mean time to system failure (MTSF) of the model is shown in figure 2. This figure indicates that MTSF decreases with the increase of failure rates λ and λ_2 for fixed values of other parameters. But, MTSF increase as repair rate θ and replacement rate β increase. Figure 3 and 4 depict the behaviour of availability and profit of the model. From these figures it can be seen that their values go on decreasing as failure rates λ and λ_2 increase. However, their values increase if repair rate θ and replacement rate β increase for fixed values of other parameters including K_1=5000, K_2=500 and K_3=50. Further, if we interchange p and q, the availability and the profit of the system increase for $\lambda \leq 0.07$.

Hence, on the basis of the results obtained for a particular case it is concluded that the concepts of priority for operation to new unit over the duplicate unit and replacement of the degraded unit at its failure are economically beneficial to use.

AVAILABILITY vs FAILURE RATE(λ) (Fig.3)



PROFIT vs FAILURE RATE(λ) (Fig.4)



References

1. Chander S. Reliability models with priority for operation and repair with arrival time of the server. Pure and Applied Mathematika Sciences 2005; LXI(1-2): 9-22.
2. Kadyan M S, Malik S C, Kumar J. Cost Analysis of a system under priority to repair and degradation. International Journal of Statistics and System 2010; 5(1): 1-10.
3. Khaled M. El-Said, Mohamed S. El-Sherbeny. Stochastic analysis of a two-unit cold standby system with two-stage repair and waiting

time. Sankhya: The Indian Journal of Statistics 2010; 72-B (1): 1-10.
4. Kumar A, Agarwal M. A review of standby redundant systems. IEEE Trans. Reliab. 1980; R-29(4): 290-294.
5. Malik S C, Chand P. Cost-Benefit analysis of a standby system with inspection subject to degradation. Aligarh Journal of Statistics 2009; 29: 25-37.
6. Malik S C, Chand P, Singh J. Stochastic analysis of an operating system with two types of inspection subject to degradation. Journal of Applied Probability and Statistics 2008; 3(2): 227-241.
7. Mokaddis G S, Labib S W, Ahmed A M. Analysis of a two-unit warm standby system subject to degradation. Microelectron. Relia. 1997; 37(4): 641-647.
8. Osaki S, Nakagawa T. Bibliography of reliability and availability of stochastic system redundant system. IEEE Trans. Reliab. 1976; R-25(4): 284-287.
9. Sridharan V, Mohanavadivu P. Stochastic behavior of a two-unit standby system with two types of repairmen and patience time. Math. Comput. Modeling 1998; 28: 63-71.
10. Wang Z, Kang R, Xie L. Dynamic reliability modeling of systems with common cause failure under random load. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2009; 3(43): 47–54.

Dr Jitender KUMAR, Ph.D., Assistant Prof.
Dr Mukender Singh KADYAN, Ph.D., Assistant Prof.
Department of Statistics & O.R.,
Kurukshetra University,
Kurukshetra (India)–136119
E-mail: khatkarjitu@gmail.com; mskadian@kuk.ac.in

Dr Suresh Chander MALIK, Ph.D., Prof.
Department of Statistics,
M. D. University, Rohtak (India)–124001
E-mail: sc_malik@rediffmail.com

Article citation info:
HUANG HZ. Structural reliability analysis using fuzzy sets theory. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 284–294.

Hong-Zhong HUANG

STRUCTURAL RELIABILITY ANALYSIS USING FUZZY SETS THEORY

ANALIZA NIEZAWODNOŚCIOWA KONSTRUKCJI Z WYKORZYSTANIEM TEORII ZBIORÓW ROZMYTYCH

Prediction of structural performance is a complex problem because of the existence of randomness and fuzziness in engineering practice. In this area, reliability analyses have been performed using probabilistic methods. This work investigates reliability analysis of structure involving fuzziness and randomness. In particular, the safety state of the structure is defined by a fuzzy state variable, fuzzy random allowable interval, or fuzzy random generalized strength. Because the membership function of the fuzzy safety state is the key to structural reliability analysis using the fuzzy sets theory, this work proposes useful methods to determine the membership functions and develops a structural reliability analysis method based on the fuzzy safety state. Several examples are provided to illustrate the proposed methods.

Keywords*: reliability, structure, fuzzy safety state, membership function, fuzzy random generalized stress, fuzzy random generalized strength.*

Przewidywanie zachowania konstrukcji stanowi złożone zagadnienie ze względu na istnienie w praktyce inżynierskiej losowości i rozmytości. Na tym obszarze, analizy niezawodnościowe prowadzono dotąd przy pomocy metod probabilistycznych. W niniejszej pracy przedstawiono metodę niezawodnościowej analizy konstrukcji uwzględniającą rozmytość i losowość. Dokładniej, stan bezpieczeństwa konstrukcji określano za pomocą rozmytej zmiennej stanu, rozmytego losowego przedziału dozwolonego lub rozmytej losowej uogólnionej wytrzymałości. Ponieważ funkcja przynależności rozmytego stanu bezpieczeństwa stanowi klucz do niezawodnościowej analizy konstrukcji wykorzystującej teorię zbiorów rozmytych, w niniejszej pracy zaproponowano przydatne metody wyznaczania funkcji przynależności oraz opracowano metodę niezawodnościowej analizy konstrukcji opartą na rozmytym stanie bezpieczeństwa. Zaproponowane metody zilustrowano kilkoma przykładami.

Słowa kluczowe*: niezawodność, konstrukcja, rozmyty stan bezpieczeństwa, funkcja przynależności, rozmyte losowe uogólnione naprężenie, rozmyta losowa uogólniona wytrzymałość.*

1. Introduction

Stress Strength Interference (SSI) is a fundamental model for structural reliability-based design and has been widely used in engineering practice [1, 6, 24, 25]. In an SSI model, a limit state function must be determined. In most studies, the limit state function is usually assumed to be exact without considering fuzziness. This means that the corresponding theory used to determine the limit states is perfect, which may not be realistic in real-world applications. Indeed, unvoidable errors could be resulted if such limit state function is used for reliability analysis. As a result, it is necessary to develop a new reliability model that takes fuzziness into consideration [27, 28]. To this end, the following question must be answered: what fundamental issue needs to be addressed for this purpose? The answer is that more data (i.e., experimental results) should be collected. As a matter of fact, if the exact value of the actual strength of a structure cannot be determined, we need to rely on more data to give additional information necessary to correct the theoretical model used [18].

When it is expensive to obtain experimental data or there are a few but poorly documented instances of failure of the prototype system, it would be difficult to correct the theoretical model. There is another extreme case where there are no data at all for calculating the probability of failure at the early design stage. For these circumstances, using engineering judgment or experience for similar structures in SSI modeling becomes a very useful alternative.

Uncertainties and ambiguities in structural performance have been dealt with using probability theory. However, it is worth pointing out that some uncertainties, which are not random in nature, may play important roles in the safety assessment of engineering structures [17]. In other words, the probability-based reliability provides a solution different from the observed failure rate which is inferred from the statistics of structural accidents [4, 28]. A more fundamental argument against the conventional approach to parameterizing model uncertainties is provided by Blockley [2, 3].

Fuzziness could be produced due to some factors, such as omissions, human error, inadequate modeling, experience, and intuition of the engineers. Such uncertainties are called "subjective uncertainties", because they could be evaluated solely by an engineer's experience and judgment. Fuzzy sets theory, which was proposed by Zadeh in 1965, is available to deal with the subjective uncertainties in a quantitative way. Moreover, this theory makes it possible to define safety events in a more flexible form than the probabilistic approach.

The first known theoretical approach to using fuzzy logic for failure diagnosis belongs to Tsukamato and Tarano [23]. Brown [4] and Blockley [3] applied the fuzzy sets theory in an attempt to explain the difference between the calculated and observed failure probabilities. Savchuk [18] suggested some improvement of reliability estimation in the framework of the SSI model, which was essentially based on a limit-state model. So far, only initial attempts have been made [14, 15], but sophisticated formulations and algorithms for numerical treatments and applications to complex structural reliability analysis have not been reported.

The approach proposed in this work aims to overcome some of the problems of the conventional treatment mentioned above. Specifically, the main purpose is to perform reliability estimation based on a limit-state model considering both the error of the limit-state model and the fuzziness of data.

2. Fuzzy safety state of structure

In a traditional SSI model, the generalized strength of a structure *R* and the generalized stress *S* are both considered to be random variables. The safety margin, also called the state variable of the structure, is defined as:

$$Z \overset{\Delta}{=} R - S \,. \tag{1}$$

It is obvious that the state variable is also a random variable. The random event that the structure works satisfactorily during its service life T, denoted by *A*, is defined as:

$$A \overset{\Delta}{=} \{S < R\} \,. \tag{2}$$

This event is also called the state of safe operation, or simply the safety state, of the structure. Let $p_S(x)$ denote the probability density function (pdf) of the generalized stress $S \in (-\infty, \infty)$, $p_R(y)$ the pdf of the generalized strength $R \in [x, \infty)$, $p_{S,R}(x,y)$ the joint pdf of *S* and *R*, and $p_Z(z)$ the pdf of the state variable Z. The reliability of the structure is simply the probability of the safety state of the structure and can be expressed by integrating the pdf of the random event *A*, with respect to the domains of *x* and *y*:

$$P_r = P(A) = \int_0^{\infty} p_Z(z)\mathrm{d}z = \int_{-\infty}^{\infty}\int_x^{\infty} p_{S,R}(x,y)\mathrm{d}y\mathrm{d}x \,. \tag{3}$$

The probability of structural failure is then given by:

$$P_f = 1 - P(A) \,. \tag{4}$$

When the random variables *S* and *R* are independent, their joint pdf can be simply expressed as the multiplication of their individual pdf:

$$p_{S,R}(x,y) = p_S(x)p_R(y) \,. \tag{5}$$

According to the SSI model, the structure is safe as long as the generalized stress *S* is lower than the generalized strength *R*. However, the accuracy of this theory has been widely questioned [13, 14, 15, 19, 21]. First, the safety criterion should be fuzzy in practice [7, 8, 9, 10, 26]. Second, the generalized strength is usually not known precisely and thus is fuzzy [11, 19]. For example, cracks had been found in the tail rotor components of a CH-149 Cormorant helicopter [22], which had been created by the generalized strength during its operations. Third, the load is also fuzzy [14, 19, 11]. In the space shuttle Columbia, the debris hitting it has led to its demise during the re-entry [13]. Moreover, a turbine blade was fractured and traveled through subsequent sections of the turbine, while a shroud that dropped into the turbine air path caused excessive wear to several turbine blades at Langley AFB [16]. Such impacts generated fuzzy random loads causing structural failures. Of course, the behavior of the structure under study and the stress developed may be fuzzy too. Solid rocket seal leakage during the launch of space shuttle Challenger was undetected and precipitated its disintegration [13]. This event has been most probability originated from the fuzzy random generalized stresses in the inter-connections. These observations have led to the fuzzy versions of SSI models.

When failure modes, such as fatigue, abrasion, and erosion, are considered, the safety state of operation of the structure under consideration may exhibit both fuzziness and randomness [21]. Therefore, both fuzziness and randomness need to be considered in reliability analysis of the structure.

In consideration of the above observations, the safety state of the structure, i.e., the state of satisfactory operation, is often treated as a fuzzy set, denoted by $\tilde{A}$, which is a subset of the universe of discourse of the state of the structure. This means that the event of the safe operation is considered to be a fuzzy event and we should use the probability of this fuzzy event to measure the reliability of the structure.

Using the equation for the probability of a fuzzy event [7, 30], we can define the reliability of the structure by multiplying the membership function of the state variable *Z*, $\mu_{\tilde{A}}(z)$, belonging to the fuzzy safety state $\tilde{A}$ with the pdf of *Z* as:

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty} \mu_{\tilde{A}}(z) p_Z(z)\mathrm{d}z \,, \tag{6}$$

Similarly, if we use the random variables *S* and *R* defined earlier, we can utilize the joint pdf of *S* and *R* to define the reliability of the structure as follows:

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mu_{\tilde{A}}(x,y) p_{S,R}(x,y)\mathrm{d}x\mathrm{d}y \,, \tag{7}$$

where $\mu_{\tilde{A}}(x,y)$ is the membership function of the fuzzy safety state $\tilde{A}$ in terms of realizations *x* and *y* of *S* and *R*, respectively. Note that we have the relationship between $\mu_{\tilde{A}}(z)$ and $\mu_{\tilde{A}}(x,y)$:

$$\mu_{\tilde{A}}(z) = \mu_{\tilde{A}}(x, z - x) \,. \tag{8}$$

Apparently, if *S* and *R* are independent, we can apply Eq. (5) to Eq. (7) and have

$$P_r = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mu_{\tilde{A}}(x,y) p_S(x) p_R(y)\mathrm{d}x\mathrm{d}y \,. \tag{9}$$

It is easy to show that the reliability of a structure reduces to its conventional reliability, if we use the following membership function for the safety state of the structure,

$$\mu_{\tilde{A}}(x,y) = \begin{cases} 1, & x < y \\ 0, & x \ge y \end{cases} \,. \tag{10}$$

As a result, Eq. (7) reduces to Eq. (3), and Eq. (9) can be written into

$$P_r = \int_{-\infty}^{+\infty} p_R(y)\left[\int_{-\infty}^{y} p_S(x)\mathrm{d}x\right]\mathrm{d}y \,. \tag{11}$$

Evidently, Eq. (11) is the same as the traditional formula for structural reliability evaluation when the stress and strength are independent. This means that the proposed model of structural reliability analysis is consistent with the model of conventional reliability analysis, and the latter is a special case of the former. The fuzzy safety state of a structure may be defined in one of three different forms.

The fuzzy safety state is defined by the state variable *Z*:

$$\tilde{A} \overset{\Delta}{=} \{Z \tilde{>} 0\}, \tag{12}$$

where $Z \tilde{>} 0$ indicates that the state variable *Z* is larger than 0 in a fuzzy sense. Here the membership function of $\tilde{A}$ is shown in Fig. 1(a). The transition curve from 0 to 1 may take a proportional, parabola, or in other forms.

The fuzzy safety state is defined by the fuzzy random generalized strength $\tilde{R}$:

$$\tilde{A} \overset{\Delta}{=} \{\tilde{S} < \tilde{R}\}, \quad \tilde{A} \overset{\Delta}{=} \{S < \tilde{R}\} \,. \tag{13}$$

The fuzzy random generalized strength $\tilde{R}$ is an index representing the generalized strength of the material of structure when both fuzziness and randomness are considered, and its membership function generally follows the shape illustrated in Fig. 1 (b).

The fuzzy safety state is defined by a fuzzy random allowable interval of the stress, $[\tilde{S}]$:

$$\tilde{A}\overset{\Delta}{=}\{\tilde{S}\subset[\tilde{S}]\},\quad \tilde{A}\overset{\Delta}{=}\{S\subset[\tilde{S}]\}, \tag{14}$$

where $\tilde{S}$ is the fuzzy random generalized stress. On the other hand, the fuzzy random allowable interval of the generalized stress reflects the fuzziness of the safety criteria used. In defining the safety criteria, one needs to consider the fuzziness of the structural responses, including the stress, deflection, deformation, frequency, etc., and the fuzziness of the allowable interval of the structural response. In other words, there is no clear boundary between what is allowed and what is not allowed. Thus, the allowable interval of the generalized stress possesses fuzziness, and its membership function generally has the shape as illustrated in Fig. 1(c). In the following sections, we discuss how to evaluate the reliability of a structure when one of the three forms of the fuzzy safety state is used.

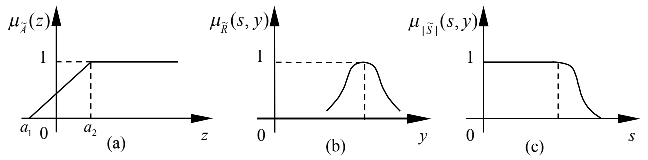


Fig. 1. Membership function

3. Form 1: The fuzzy safety state defined by the state variable

In this section, we provide a method to analyze the reliability of a structure when the fuzzy safety state is defined by the state variable. In this form, the reliability is computed simply by integrating the membership function of $\tilde{A}$ times the pdf of *Z*. Under this definition, we consider two shapes of the membership function of the fuzzy safety state, namely, the rising half-trapezoidal distribution and the rising half-ridge distribution. When the shape of the membership function is specified, we consider the commonly used pdf of exponential distribution, normal distribution, lognormal distribution, or Weibull distribution.

3.1. The membership function of the fuzzy safety state follows a rising half-trapezoidal distribution

In this case, the membership function of the fuzzy safety state $\tilde{A}$ is illustrated in Fig.1(a), and its mathematical form is given below, which represents a proportional type transition,

$$\mu_{\tilde{A}}(z)=\begin{cases}0 & ,\ z<a_1\\ \dfrac{z-a_1}{a_2-a_1} & ,\ a_1\le z\le a_2\\ 1 & ,\ z>a_2\end{cases} \tag{15}$$

3.1.1. The state variable follows the exponential distribution

In this case, the pdf of *Z* is exponential function with failure rate λ

$$p_Z(z)=\lambda e^{-\lambda z}\,. \tag{16}$$

With Eq. (6), we have the following expression for the reliability of the structure:

$$\begin{aligned}P_r&=P(\tilde{A})=\int_{-\infty}^{+\infty}\mu_{\tilde{A}}(z)p_Z(z)\mathrm{d}z\\&=\int_{a_1}^{a_2}\frac{z-a_1}{a_2-a_1}\lambda e^{-\lambda z}\mathrm{d}z+\int_{a_2}^{+\infty}\lambda e^{-\lambda z}\mathrm{d}z\\&=\frac{1}{\lambda(a_2-a_1)}\left(e^{-\lambda a_1}-e^{-\lambda a_2}\right)\end{aligned} \tag{17}$$

3.1.2. The state variable follows a normal distribution

In this case, the pdf of *Z* is expressed as:

$$p_Z(z)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(z-\mu)^2}{2\sigma^2}}, \tag{18}$$

where μ and σ^2 are the mean and variance of Z , respectively. With Eq. (6), the reliability of the structure can be expressed as:

$$\begin{aligned}P_r&=P(\tilde{A})=\int_{-\infty}^{+\infty}\mu_{\tilde{A}}(z)p_Z(z)\mathrm{d}z\\&=\int_{a_1}^{a_2}\frac{z-a_1}{a_2-a_1}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(z-\mu)^2}{2\sigma^2}}\mathrm{d}z+\int_{a_2}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(z-\mu)^2}{2\sigma^2}}\mathrm{d}z\\&=1+\left(\frac{\mu-a_2}{a_2-a_1}\right)\Phi\left(\frac{a_2-\mu}{\sigma}\right)-\left(\frac{\mu-a_1}{a_2-a_1}\right)\Phi\left(\frac{a_1-\mu}{\sigma}\right)\\&\quad+\frac{\sigma}{(a_2-a_1)\sqrt{2\pi}}\left(e^{-\frac{(a_1-\mu)^2}{2\sigma^2}}-e^{-\frac{(a_2-\mu)^2}{2\sigma^2}}\right)\end{aligned} \tag{19}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of a standard normal random variable.

3.1.3. The state variable follows a lognormal distribution

Herein, the pdf of *Z* is written as the lognormal form

$$p_Z(z)=\frac{1}{\sqrt{2\pi}z\sigma_{\ln z}}e^{-\frac{(\ln z-\mu_{\ln z})^2}{2\sigma_{\ln z}^2}}, \tag{20}$$

where $\mu_{\ln z}$ and $\sigma_{\ln z}^2$ are the mean and variance of $\ln Z$. Then the reliability can be obtained by

$$\begin{aligned}P_r&=P(\tilde{A})=\int_{-\infty}^{+\infty}\mu_{\tilde{A}}(z)p_Z(z)\mathrm{d}z\\&=\int_{a_1}^{a_2}\frac{z-a_1}{a_2-a_1}\frac{1}{\sqrt{2\pi}z\sigma_{\ln z}}e^{-\frac{(\ln z-\mu_{\ln z})^2}{2\sigma_{\ln z}^2}}\mathrm{d}z+\int_{a_2}^{+\infty}\frac{1}{\sqrt{2\pi}z\sigma_{\ln z}}e^{-\frac{(\ln z-\mu_{\ln z})^2}{2\sigma_{\ln z}^2}}\mathrm{d}z\\&=1+e^{\mu_{\ln z}+\frac{1}{2}\sigma_{\ln z}^2}\left[\Phi\left(\frac{\ln a_2-\mu_{\ln z}-\sigma_{\ln z}^2}{\sigma_{\ln z}}\right)^2-\Phi\left(\frac{\ln a_1-\mu_{\ln z}-\sigma_{\ln z}^2}{\sigma_{\ln z}}\right)\right]\\&\quad+\frac{a_1}{a_2-a_1}\Phi\left(\frac{\ln a_1-\mu_{\ln z}}{\sigma_{\ln z}}\right)-\frac{a_2}{a_2-a_1}\Phi\left(\frac{\ln a_2-\mu_{\ln z}}{\sigma_{\ln z}}\right)\end{aligned} \tag{21}$$

3.1.4. The state variable follows the 3-parameter Weibull distribution

In this case, the pdf of *Z* is

$$p_Z(z)=\frac{\beta}{\eta}\left(\frac{z-\gamma}{\eta}\right)^{\beta-1}e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}}. \tag{22}$$

Then the reliability of the structure can be expressed as:

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty} \mu_{\tilde{A}}(z) p_Z(z) \mathrm{d}z$$

$$= \int_{a_1}^{a_2} \frac{z-a_1}{a_2-a_1} \frac{\beta}{\eta}\left(\frac{z-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}} \mathrm{d}z + \int_{a_2}^{+\infty} \frac{\beta}{\eta}\left(\frac{z-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}} \mathrm{d}z$$

$$= \frac{1}{a_2-a_1} \int_{a_1}^{a_2} e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}} \mathrm{d}z \,, \tag{23}$$

where the integral can be solved numerically.

3.2. The membership function of the fuzzy safety state follows a rising half-ridge distribution

In this case, the membership function of the fuzzy safety state $\tilde{A}$ is illustrated in Fig. 2, where the incremental process is in a sinusoidal form given by:

$$\mu_{\tilde{A}}(z) = \begin{cases} 0 & , z < a_1 \\ \frac{1}{2} + \frac{1}{2}\sin\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right) & , a_1 \le z \le a_2 \\ 1 & , z > a_2 \end{cases} \tag{24}$$

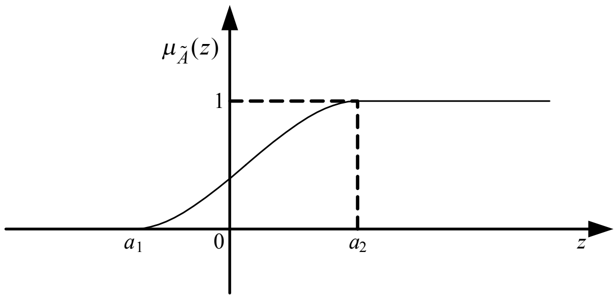


Fig. 2. Rising half-ridge distribution

3.2.1. The state variable follows an exponential distribution

The pdf of Z is

$p_Z(z) = \lambda e^{-\lambda z}$

Then the reliability of the structure can be expressed as:

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty} \mu_{\tilde{A}}(z) p_Z(z) \mathrm{d}z$$

$$= \int_{a_1}^{a_2} \left[\frac{1}{2} + \frac{1}{2}\sin\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right)\right] \lambda e^{-\lambda z} \mathrm{d}z + \int_{a_2}^{+\infty} \lambda e^{-\lambda z} \mathrm{d}z$$

$$= \frac{\pi^2}{2(\lambda^2 (a_2-a_1)^2 + \pi^2)}\left(e^{-\lambda a_1} + e^{-\lambda a_2}\right) \tag{25}$$

3.2.2. The state variable follows a normal distribution

In this case, the pdf of Z is given by

$$p_Z(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

Then the reliability of the structure is

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty} \mu_{\tilde{A}}(z) p_Z(z) \mathrm{d}z$$

$$= \int_{a_1}^{a_2} \left[\frac{1}{2} + \frac{1}{2}\sin\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right)\right] \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \mathrm{d}z + \int_{a_2}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \mathrm{d}z$$

$$= 1 - \frac{1}{2}\Phi\left(\frac{a_1-\mu}{\sigma}\right) - \frac{1}{2}\Phi\left(\frac{a_2-\mu}{\sigma}\right)$$

$$+ \frac{1}{2\sqrt{2\pi}\sigma} \int_{a_1}^{a_2} \sin\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right) e^{-\frac{(z-\mu)^2}{2\sigma^2}} \mathrm{d}z \,, \tag{26}$$

where the integral can be solved by a numerical method.

3.2.3. The state variable follows a lognormal distribution

In this case, the pdf of Z is expressed as:

$$p_Z(z) = \frac{1}{\sqrt{2\pi} z \sigma_{\ln z}} e^{-\frac{(\ln z - \mu_{\ln z})^2}{2\sigma_{\ln z}^2}}$$

Then the reliability of the structure is written as:

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty} \mu_{\tilde{A}}(z) p_Z(z) \mathrm{d}z$$

$$= \int_{a_1}^{a_2} \left[\frac{1}{2} + \frac{1}{2}\sin\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right)\right] \frac{1}{\sqrt{2\pi} z \sigma_{\ln z}} e^{-\frac{(\ln z - \mu_{\ln z})^2}{2\sigma_{\ln z}^2}} \mathrm{d}z$$

$$+ \int_{a_2}^{+\infty} \frac{1}{\sqrt{2\pi} z \sigma_{\ln z}} e^{-\frac{(\ln z - \mu_{\ln z})^2}{2\sigma_{\ln z}^2}} \mathrm{d}z$$

$$= 1 - \frac{1}{2}\Phi\left(\frac{\ln a_1 - \mu_{\ln z}}{\sigma_{\ln z}}\right) - \frac{1}{2}\Phi\left(\frac{\ln a_2 - \mu_{\ln z}}{\sigma_{\ln z}}\right)$$

$$+ \frac{1}{2\sqrt{2\pi}\sigma_{\ln z}} \int_{a_1}^{a_2} \frac{1}{z} e^{-\frac{(\ln z - \mu_{\ln z})^2}{2\sigma_{\ln z}^2}} \sin\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right) \mathrm{d}z \,, \tag{27}$$

where the integral can be solved by a numerical method.

3.2.4. The state variable follows a Weibull distribution

The pdf of Z is written as

$$p_Z(z) = \frac{\beta}{\eta}\left(\frac{z-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}} .$$

Then the reliability of the structure is given by

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty} \mu_{\tilde{A}}(z) p_Z(z) \mathrm{d}z \tag{28}$$

$$= \int_{a_1}^{a_2} \left[\frac{1}{2} + \frac{1}{2}\sin\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right)\right] \frac{\beta}{\eta}\left(\frac{z-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}} \mathrm{d}z + \int_{a_2}^{+\infty} \frac{\beta}{\eta}\left(\frac{z-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}} \mathrm{d}z$$

$$= \frac{1}{2}\left(e^{-\left(\frac{a_1-\gamma}{\eta}\right)^{\beta}} + e^{-\left(\frac{a_2-\gamma}{\eta}\right)^{\beta}}\right) + \frac{\beta}{2\eta} \int_{a_1}^{a_2} \sin\left[\frac{\pi}{a_2-a_1}\left(z - \frac{a_1+a_2}{2}\right)\right] \left(\frac{z-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{z-\gamma}{\eta}\right)^{\beta}} \mathrm{d}z \,,$$

where the integral can be solved via a numerical method.

3.3. Simply Supported Beam under Stress

Consider a beam which is supported at both ends carrying uniformly distributed load as shown in Fig. 3. All concerned variables, including the dimensions of the beam, the load distribution, and the strength of the material, are assumed to follow the normal distributions,

$$q \sim \mathrm{N}(210, 7^2) \text{ N/mm},$$

$$b \sim \mathrm{N}(120, 10^2) \text{ mm},$$

$$l \sim \mathrm{N}(4000,150^2)\ \mathrm{mm},$$

$$h \sim \mathrm{N}(240,10^2)\ \mathrm{mm},$$

$$R \sim \mathrm{N}(623,23^2)\ \mathrm{MPa}.$$

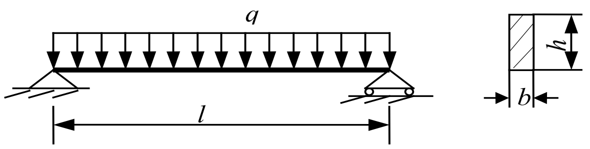


Fig. 3. The simply supported beam

We first apply the form 1 model where the fuzzy safety state is defined by the state variable *Z*. The membership function of the fuzzy safety state $\tilde{A}$ is given by the following half-trapezoidal distribution:

$$\mu_{\tilde{A}}(z)=\begin{cases}0 & ,z<-40\\ \dfrac{z+40}{80} & ,-40\le z\le 40\\ 1 & ,z>40\end{cases}$$

Then the reliability of this beam can be calculated as follows.

The maximum stress in the simple supported beam is given by

$$S=\frac{0.75ql^2}{bh^2}.$$

Since q,l,b,h are all normal random variables, we can obtain the mean and the approximate standard deviation using the Taylor expansion of the maximum stress S with the following equations:

$$\mu_S=\frac{0.75\mu_q\mu_l^2}{\mu_b\mu_h^2}=\frac{0.75\times 210\times 4000^2}{120\times 240^2}=364.6\mathrm{MPa}$$

$$\sigma_S\overset{\Delta}{=}\left\{\left[\frac{\partial S}{\partial q}\Big|_{X=\mu_X}\right]^2\sigma_q^2+\left[\frac{\partial S}{\partial l}\Big|_{X=\mu_X}\right]^2\sigma_l^2+\left[\frac{\partial S}{\partial b}\Big|_{X=\mu_X}\right]^2\sigma_b^2+\left[\frac{\partial S}{\partial h}\Big|_{X=\mu_X}\right]^2\sigma_h^2\right\}^{\frac{1}{2}}$$

= 52.4 MPa,

where $X=(X_1,X_2,X_3,X_4)=(q,l,b,h)$ and $\mu_X=(\mu_q,\mu_l,\mu_b,\mu_h)$.

The mean and the standard deviation of the state variable are

$$\mu_Z=\mu_R-\mu_S=623-364.6=258.4\mathrm{MPa}$$

$$\sigma_Z=\left(\sigma_R^2+\sigma_S^2\right)^{\frac{1}{2}}=(23^2+52.4^2)^{\frac{1}{2}}=57.2\mathrm{MPa}$$

The pdf of Z can be considered to follow the normal distribution

$$p_Z(z)=\frac{1}{\sqrt{2\pi}\sigma_Z}e^{-\frac{(z-\mu_Z)^2}{2\sigma_Z^2}}=\frac{1}{\sqrt{2\pi}\times 57.2}e^{-\frac{(z-258.4)^2}{2\times 57.2^2}}$$

Using the obtained $\mu_{\tilde{A}}(z)$, $p_Z(z)$, and Eq. (6), we have $P_r\approx 0.99928$. If we use the conventional reliability analysis method, the reliability of the beam is

$$P_r=\Phi(\frac{\mu_Z}{\sigma_Z})=\Phi(4.517)=0.99999.$$

From this example, we see that the reliability of the beam obtained with the conventional reliability analysis method is higher than that with the proposed method. The conventional reliability method over-estimates the reliability of the beam, and thus the obtained design is risker than that obtained through the proposed method. This is because the conventional reliability method does not consider the fuzziness of the safety criterion.

4. Form 2: The fuzzy safety state defined by the fuzzy random generalized strength

In this form, the reliability is computed by integrating the membership functions of $\tilde{A}$ which is the weighted sum of $\mu_{\tilde{A}_1}(x,y)$ and $\mu_{\tilde{A}_2}(x,y)$. $\mu_{\tilde{A}_1}(x,y)$ is the area ratio of $\mu_{\tilde{R}}(s,y)$, while $\mu_{\tilde{A}_2}(x,y)$ is the area ratio of $\mu_{\tilde{R}_{\mathrm{peak}}}(s,y)$ in the case of random generalized stress and fuzzy random generalized strength. The reliability is computed by integrating the membership functions of $\tilde{A}$, which equals the weighted sum of $\mu_{\tilde{A}_1'}(x,y)$ and $\mu_{\tilde{A}_2'}(x,y)$. $\mu_{\tilde{A}_1'}(x,y)$ is the area ratio between the difference of $\mu_{\tilde{S}}(s,x)$, $\mu_{\tilde{R}}(s,y)$ and $\mu_{\tilde{S}}(s,x)$, while $\mu_{\tilde{A}_2'}(x,y)$ is the membership function of $\mu_{\tilde{R}}(S_{\mathrm{peak}}(x),y)$ in the case of fuzzy random generalized stress and fuzzy random generalized strength.

4.1. Two characteristic values of a fuzzy number

A real fuzzy number $\tilde{N}$ is defined as a fuzzy set in the domain of real numbers $\mathbb{R}$, and its membership function (shown in Fig. 4) has the following characteristics:

(1) It is a continuous mapping from $\mathbb{R}$ to the closed interval $[0,w]$, $0<\mathrm{w}\le 1$;

(2) It is equal to 0 in $(-\infty,c]$, i.e., $\mu_{\tilde{N}}(s)=0$, when $-\infty<s\le c$;

(3) It is strictly increasing in $[c,a]$;

(4) It is equal to 1 in $[a,b]$, i.e., $\mu_{\tilde{N}}(s)=1$, when $a\le s\le b$;

(5) It is strictly decreasing in $[b,d]$;

(6) It is equal to 0 in $[d,\infty)$, i.e. $\mu_{\tilde{N}}(s)=0$, when $d\le s<\infty$;

where *a*, *b*, *c* and *d* are real numbers and $c\le a\le b\le d$. Among other choices, we may set $c=-\infty$, $a=b$, $c=a$, $b=d$, and $d=+\infty$ individually or in various combinations. When the generalized strength and the generalized stress in structural safety analysis are treated as fuzzy variables, their membership functions often exhibit the shape depicted in Fig. 5. From detailed analysis of the membership function of a fuzzy variable, one can see that it can be well represented by two characteristic values, one is the area distribution of the membership function and the other is the position of the peak value. The area distribution of a fuzzy variable, i.e., to a certain point s, the area on its left and right sides under the membership function curve, is analogous to the pdf of a random variable. The peak value of a fuzzy number shown in Fig. 4 is

$$s_{\mathrm{peak}}=(a+b)/2 \tag{29}$$

4.2. Determination of the membership function of fuzzy safety state in the case of random generalized stress and fuzzy random generalized strength

The fuzzy safety event in the case of random generalized stress and fuzzy random generalized strength is defined in Eq. (13), with fuzzy safety state in this case being denoted by $\tilde{A}$. According to the area distribution of the membership function $\mu_{\tilde{R}}(s,y)$ of fuzzy random generalized strength $\tilde{R}$ and the position of the peak value $R_{\mathrm{peak}}(y)$ of $\mu_{\tilde{R}}(s,y)$, two partial expressions of $\mu_{\tilde{A}}(x,y)$, $\mu_{\tilde{A}_1}(x,y)$ and $\mu_{\tilde{A}_2}(x,y)$, can be obtained. Multiplying these two factors by weights w_1 and w_2 respectively (w_1 + w_2 = 1), $\mu_{\tilde{A}}(x,y)$ becomes the weighted sum of these two partial expressions:

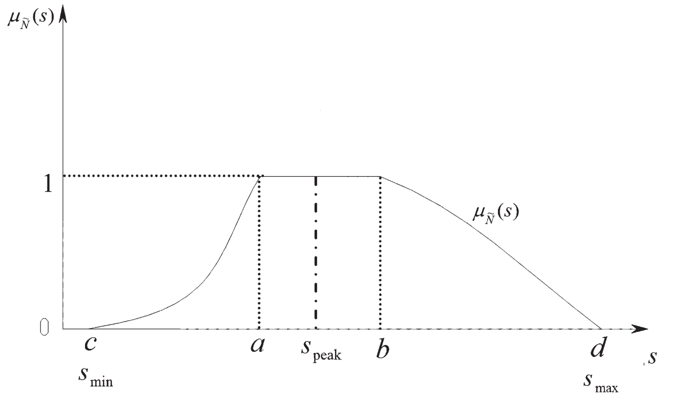


Fig. 4. The membership function of a fuzzy number

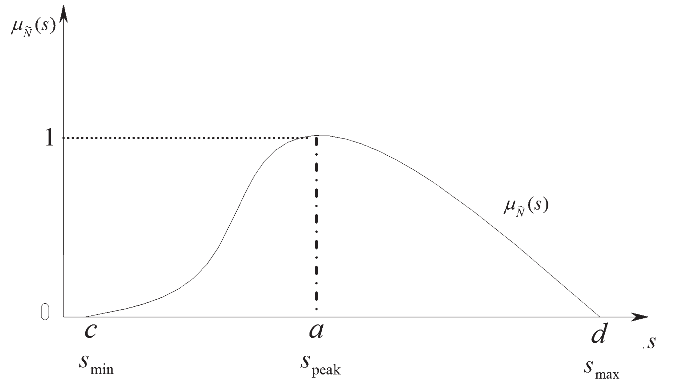


Fig. 5. The membership function of generalized strength and generalized stress

$$\mu_{\tilde{A}}(x,y) = w_1\mu_{\tilde{A}_1}(x,y) + w_2\mu_{\tilde{A}_2}(x,y) \tag{30}$$

4.2.1. Determination of $\mu_{\tilde{A}_1}(x,y)$

As shown in Fig. 6, according to the area distribution of the membership function $\mu_{\tilde{R}}(s,y)$ of fuzzy random generalized strength $\tilde{R}$, the partial expression $\mu_{\tilde{A}_1}(x,y)$ of $\mu_{\tilde{A}}(x,y)$ has the expression of integral value of $\mu_{\tilde{R}}(s,y)$ in the domain $\left[s(x), R_{\max}(y)\right]$ divided by the total integral value of $\mu_{\tilde{R}}(s,y)$ in the full domain $\left[R_{\min}(y), R_{\max}(y)\right]$

$$\mu_{\tilde{A}_1}(x,y) = \frac{\int_{s(x)}^{R_{\max}(y)} \mu_{\tilde{R}}(s,y)\mathrm{d}s}{\int_{R_{\min}(y)}^{R_{\max}(y)} \mu_{\tilde{R}}(s,y)\mathrm{d}s} \tag{31}$$

4.2.2. Determination of $\mu_{\tilde{A}_2}(x,y)$

According to the relative position between generalized stress $s(x)$ and $R_{\text{peak}}(y)$, the partial expression $\mu_{\tilde{A}_2}(x,y)$ of $\mu_{\tilde{A}}(x,y)$ can be determined. If $R_{\text{peak}}(y)$ is considered to be a deterministic value, $\mu_{\tilde{A}_2}(x,y)$ varies from 1 to 0 at $R_{\text{peak}}(y)$ when $s(x)$ passes through $R_{\text{peak}}(y)$ from left to right, which leads to the discontinuity of the membership function $\mu_{\tilde{A}}(x,y)$ at $R_{\text{peak}}(y)$. To overcome this problem, $R_{\text{peak}}(y)$ is converted to a fuzzy set $\tilde{R}_{\text{peak}}$, whose membership function $\mu_{\tilde{R}_{\text{peak}}}(s,y)$ is a normal or symmetric triangular membership function. The range δ_2 of $\tilde{R}_{\text{peak}}$ is determined according to how steep the change of $\mu_{\tilde{A}_2}(x,y)$ should be near $R_{\text{peak}}(y)$: the steeper the change of $\mu_{\tilde{A}_2}(x,y)$ is, the smaller δ_2 is. As shown in Fig. 7, $\mu_{\tilde{A}_2}(x,y)$ can be determined as the ratio between the integral of $\mu_{\tilde{R}_{\text{peak}}}(s,y)$ in the domain $\left[s(x), R_{\text{peak}}(y)+\delta_2\right]$ and the integral of $\mu_{\tilde{R}_{\text{peak}}}(s,y)$ in the domain $\left[R_{\text{peak}}(y)-\delta_2, R_{\text{peak}}(y)+\delta_2\right]$ as

$$\mu_{\tilde{A}_2}(x,y) = \frac{\int_{s(x)}^{R_{\min}(y)+\delta_1+\delta_2} \mu_{\tilde{R}_{\text{peak}}}(s,y)\mathrm{d}s}{\int_{R_{\min}(y)+\delta_1-\delta_2}^{R_{\min}(y)+\delta_1+\delta_2} \mu_{\tilde{R}_{\text{peak}}}(s,y)\mathrm{d}s}, \tag{32}$$

where $R_{\text{peak}}(y) = R_{\min}(y) + \delta_1$.

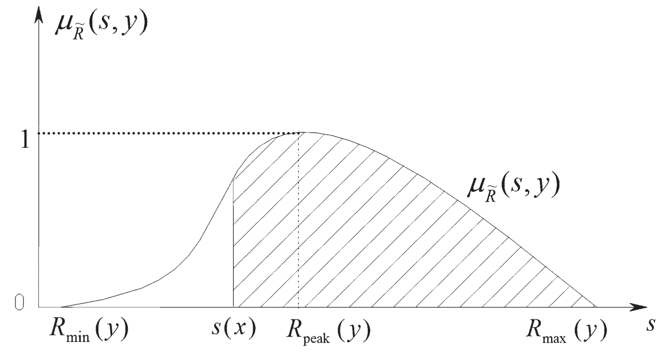


Fig. 6. Determine $\mu_{\tilde{A}_1}(x,y)$ according to the area distribution of $\mu_{\tilde{R}}(s,y)$

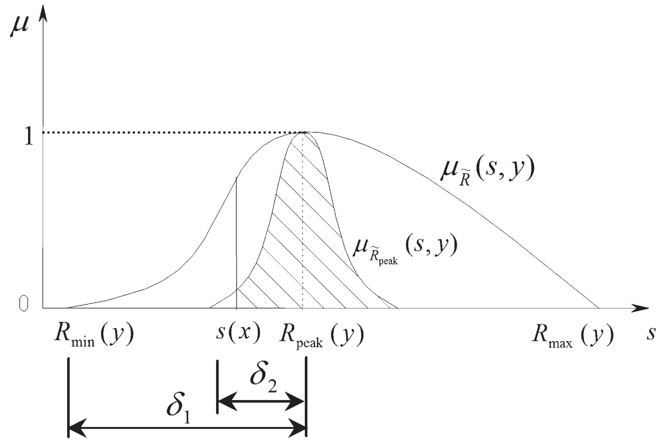


Fig. 7. Determine $\mu_{\tilde{A}_2}(x,y)$ according to the position of $R_{\text{peak}}(y)$

4.2.3. Determination of weights w_1 and w_2

The weights w_1 and w_2 denote the relative influence of the corresponding characteristic factors on $\mu_{\tilde{A}}(x,y)$. w_1 and w_2 can be determined based on experience and through other methods that may be problem-specific. Generally, the larger is the dissymmetry of the membership function $\mu_{\tilde{R}}(s,y)$, the larger is w_2. Based on Eqs. (30-32), the membership function of fuzzy safety state in the case of random generalized stress and fuzzy random generalized strength, $\mu_{\tilde{A}}(x,y)$, takes the form

$$\mu_{\tilde{A}}(x,y)=w_1\frac{\int_{s(x)}^{R_{\max}(y)}\mu_{\tilde{R}}(s,y)\mathrm{d}s}{\int_{R_{\min}(y)}^{R_{\max}(y)}\mu_{\tilde{R}}(s,y)\mathrm{d}s}+w_2\frac{\int_{s(x)}^{R_{\min}(y)+\delta_1+\delta_2}\mu_{\tilde{R}_{\mathrm{peak}}}(s,y)\mathrm{d}s}{\int_{R_{\min}(y)+\delta_1-\delta_2}^{R_{\min}(y)+\delta_1+\delta_2}\mu_{\tilde{R}_{\mathrm{peak}}}(s,y)\mathrm{d}s}\,. \quad (33)$$

Therefore the reliability can be computed using this membership function as

$$P_r=P(\tilde{A})=\int_{R_{\min}}^{R_{\max}}\int_{S_{\min}}^{S_{\max}}\mu_{\tilde{A}}(x,y)p_S(x)p_R(y)\mathrm{d}x\mathrm{d}y\,, \quad (34)$$

where $p_S(x)$ is the pdf of the random generalized stress and $p_R(y)$ is the pdf of the fuzzy random generalized strength.

4.3. Determination of the membership function of fuzzy safety state in the case of fuzzy random generalized stress and fuzzy random generalized strength

Fuzzy safety state in the case of fuzzy random generalized stress and fuzzy random generalized strength is denoted by $\tilde{A}'$, and the corresponding membership function is $\mu_{\tilde{A}'}(x,y)$. Based on the membership function of fuzzy safety state $\tilde{A}'$, $\mu_{\tilde{A}}(x,y)$, the area distribution of the membership function of fuzzy random generalized stress $\tilde{S}$, $\mu_{\tilde{S}}(s,x)$, and position of the peak value of $\mu_{\tilde{S}}(s,x)$, $S_{\mathrm{peak}}(x)$, two partial expressions of $\mu_{\tilde{A}'}(x,y)$, $\mu_{\tilde{A}_1'}(x,y)$ and $\mu_{\tilde{A}_2'}(x,y)$ can be obtained. Multiplying these two factors by weights w_1' and w_2' respectively ($w_1'+w_2'$=1), and $\mu_{\tilde{A}'}(x,y)$ equals to the weighted sum of the two partial expressions:

$$\mu_{\tilde{A}'}(x,y)=w_1'\mu_{\tilde{A}_1'}(x,y)+w_2'\mu_{\tilde{A}_2'}(x,y) \quad (35)$$

Agin, the weights w_1' and w_2' denote the relative influence of the corresponding characteristic factors on $\mu_{\tilde{A}'}(x,y)$. w_1' and w_2' can be determined based on experience and through other problem-specific methods. Generally, the larger is the asymmetry of the membership function $\mu_{\tilde{S}}(s,x)$, the larger is w_2'.

4.3.1. Determination of $\mu_{\tilde{A}_1'}(x,y)$

As shown in Fig. 8, according to the area distribution of the membership function of fuzzy random generalized stress $\tilde{S}$, $\mu_{\tilde{S}}(s,x)$, the partial expression of $\mu_{\tilde{A}'}(x,y)$, $\mu_{\tilde{A}_1'}(x,y)$ is the ratio of between the integral of $\mu_{\tilde{S}}(s,x)\mu_{\tilde{R}}(s,y)$ and the integral of $\mu_{\tilde{S}}(s,x)$ in the domain of $\left[s,S_{\max}(y)\right]$

$$\mu_{\tilde{A}_1'}(x,y)=\frac{\int_{s}^{S_{\max}}\mu_{\tilde{S}}(s,x)\mu_{\tilde{R}}(s,y)\mathrm{d}s}{\int_{S_{\min}}^{S_{\max}}\mu_{\tilde{S}}(s,x)\mathrm{d}s}\,. \quad (36)$$

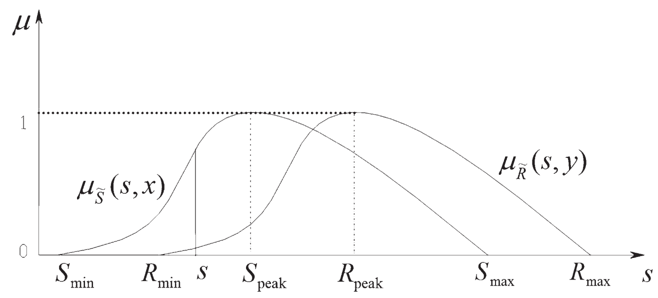


Fig. 8. Membership function of fuzzy random generalized stress and fuzzy random generalized strength

For a certain element s of fuzzy set $\tilde{S}$, the membership function value of s in fuzzy safety state $\tilde{A}'$ is $\mu_{\tilde{A}'}(s,y)$, and the membership function value of s in $\tilde{S}$ is $\mu_{\tilde{S}}(s,x)$, which can be considered to be the weight. Thus, the method to determine the expression $\mu_{\tilde{A}_1'}(x,y)$ is essentially a weighted-average method, as shown in Eq. (36).

4.3.2. Determination of $\mu_{\tilde{A}_2'}(x,y)$

According to the relative position between $S_{\mathrm{peak}}(x)$ and $\tilde{R}$, $\mu_{\tilde{A}_2'}(x,y)$ takes the expression

$$\mu_{\tilde{A}_2'}(x,y)=\mu_{\tilde{R}}(S_{\mathrm{peak}}(x),y) \quad (37)$$

On the basis of Eqs. (35–37), the membership function of fuzzy safety state in the case of fuzzy random generalized stress and fuzzy random generalized strength, $\mu_{\tilde{A}'}(x,y)$, takes the form

$$\mu_{\tilde{A}'}(x,y)=w_1'\frac{\int_{S_{\min}}^{S_{\max}}\mu_{\tilde{S}}(s,x)\mu_{\tilde{R}}(s,y)\mathrm{d}s}{\int_{S_{\min}}^{S_{\max}}\mu_{\tilde{S}}(s,x)\mathrm{d}s}+w_2'\mu_{\tilde{R}}(S_{\mathrm{peak}}(x),y) \quad (38)$$

4.4. Reliability of output axis of gearbox

In engineering, the overall structure of a gearbox is complex, which makes it more difficult to analyze the stress-strength relationship. For a gearbox, Form 1 is too simple and does not comply with this structure. In many applications, gearboxes are usually damaged by catastrophic loads such as impactions which are directly related to the strength. Therefore, damage may not be related to stress only, and form 3 may not be applicable in this situation. For simplicity, we consider a fuzzy reliability computation problem involving fuzzy random stress and fuzzy strength using the form 2 where the fuzzy safety state is defined by the fuzzy random generalized strength. It is known that the strength $\tilde{R}$ of the output axis of some gearbox is near 240MPa. As shown in Fig. 9, the membership function of $\tilde{R}$ is

$$\mu_{\tilde{R}}(s)=\begin{cases}0\,, & s\le 220\\ (s-220)/20, & 220<s\le 240\\ (280-s)/40, & 240<s\le 280\\ 0\,, & s>280\end{cases} \quad (39)$$

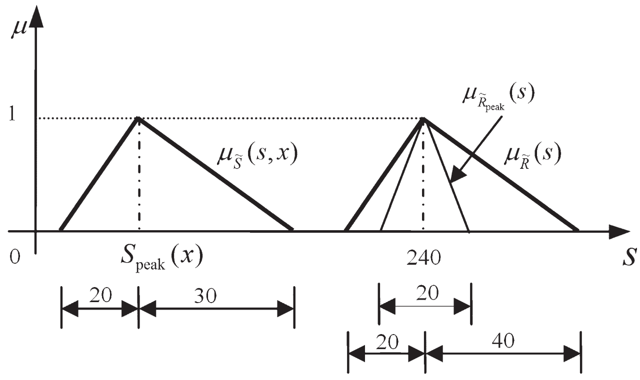


Fig. 9. The membership function of fuzzy random stress and fuzzy strength

The peak value of fuzzy random stress $\tilde{S}$, $S_{\mathrm{peak}}(x)$, follows the normal distribution $\sigma\sim \mathrm{N}(140,20^2)$ MPa. The membership function of $\tilde{S}$ is

$$\mu_{\tilde{S}}(s,S_{\text{peak}}) = \begin{cases} 0 & , \quad s \le (S_{\text{peak}} - 20) \\ (s - S_{\text{peak}} + 20)/20, & (S_{\text{peak}} - 20) < s \le S_{\text{peak}} \\ (S_{\text{peak}} + 30 - s)/30, & S_{\text{peak}} < s \le (S_{\text{peak}} + 30) \\ 0 & , \quad s > (S_{\text{peak}} + 30) \end{cases} \tag{40}$$

In the domain $[S_{\min}, S_{\max}]$, where $S_{\min} = S_{\text{peak}}(x) - 20$ and $S_{\max} = S_{\text{peak}}(x) + 30$, we now calculate the fuzzy reliability of the output axis. According to the asymmetry conditions of the membership functions of the fuzzy random stress and the fuzzy strength, the weight values are shown in Table 1. The membership functions of fuzzy random stress and fuzzy strength are both sectional functions, and it is difficult to obtain the analytical expressions for the membership function of fuzzy safety state and the fuzzy reliability. To overcome this technical difficulty, numerical methods are adopted, and MATLAB is used in implementing the computation of the membership function of fuzzy safety state and the fuzzy reliability.

Table 1. The weight values of partial expressions`

w_1	w_2	w_1'	w_2'
0.7	0.3	0.8	0.2

With different peak values of fuzzy random stress $\tilde{S}$, S_{peak}, the corresponding membership function values of fuzzy safety state are shown in Table 2. When $\tilde{S}$ is completely on the left-hand side of $\tilde{R}$, $\mu_{\tilde{A}'}(S_{\text{peak}})$ equals 1. When $\tilde{S}$ is completely on the right-hand side of $\tilde{R}$, $\mu_{\tilde{A}'}(S_{\text{peak}})$ equals 0. When $\tilde{S}$ moves through $\tilde{R}$ from the left to the right, $\mu_{\tilde{A}'}(S_{\text{peak}})$ decreases from 1 to 0 continuously and monotonically. Such a computational result is reasonable. Based on the known conditions, the pdf of the peak value S_{peak} of the fuzzy random stress is

$$p(S_{\text{peak}}) = \frac{1}{\sqrt{2\pi} \times 20} \exp\left(-\frac{(S_{\text{peak}} - 140)^2}{2 \times 20^2} \right) \tag{41}$$

From Eq. (6), the fuzzy reliability of the output axis of the reducing gearbox is

$$P_r = P(\tilde{A}') = \int_{S_{\text{peak}}(x)-20}^{S_{\text{peak}}(x)+30} \mu_{\tilde{A}'}(S_{\text{peak}}) p(S_{\text{peak}}) \mathrm{d}S_{\text{peak}} = 0.99998869 \tag{42}$$

5. Form 3: The fuzzy safety state defined by fuzzy random allowable interval

When the generalized stress is a random variable and the allowable interval of generalized stress is a fuzzy random variable, fuzzy event $\tilde{A}$ is a special fuzzy event. When both the generalized stress and its allowable interval are fuzzy random variables, $\tilde{A}$ is a general fuzzy event.

5.1. Membership function of the special fuzzy event

If any realization x of the random generalized stress is within the interval of $\mu_{[\tilde{S}]}(s,y) = 1$ where the domain of $[\tilde{S}]$ is $\left[[\tilde{S}]_{\min}, [\tilde{S}]_{\max}\right]$ as shown in Fig.1 (c), the structure is absolutely safe, thus $\mu_{\tilde{A}}(x,y) = 1$. When x is within the interval of transition, that is, $0 < \mu_{[\tilde{S}]}(s,y) < 1$, the structure is safe to some extent (depending on the value of $\mu_{[\tilde{S}]}(s,y)$). When x is completely out of $\mu_{[\tilde{S}]}(s,y)$, that is, $\mu_{[\tilde{S}]}(s,y) = 0$, the structure will absolutely fail. Therefore the membership function of the special fuzzy event $\tilde{A}$ can be defined as

$$\mu_{\tilde{A}}(x,y) \overset{\Delta}{=} \mu_{[\tilde{S}]}(s,y)|_{s=x} = \mu_{[\tilde{S}]}(x,y) \tag{43}$$

When only the randomness of the generalized stress and fuzziness of the allowable interval are taken into account, $\mu_{[\tilde{S}]}(s,y)$ degenerates to be $\mu_{[\tilde{S}]}(s)$ and

$$\mu_{\tilde{A}}(x) \overset{\Delta}{=} \mu_{[\tilde{S}]}(s)|_{s=x} = \mu_{[\tilde{S}]}(x)$$

5.2. Membership function of the general fuzzy event

As for the general fuzzy event, the degree of structural safety also depends on the relative position of membership functions $\mu_{\tilde{S}}(s,x)$ and $\mu_{[\tilde{S}]}(s,y)$. By imitating Eq. (43), the membership function of the general fuzzy event is defined as

$$\mu_{\tilde{A}}(x,y) = \mu_{[\tilde{S}]}(s,y)|_{s=\tilde{S}} = \mu_{[\tilde{S}]}(\tilde{S},y). \tag{44}$$

It is known that if $\mu_{\tilde{S}}(s,x)$ is completely covered by $\mu_{[\tilde{S}]}(s,y)$ the structure is thought to be absolutely safe, i.e., $\mu_{\tilde{A}}(x,y) = 1$. In other words, if any generalized stress $s \in \tilde{S}$ satisfies

$\mu_{[\tilde{S}]}(s,y) \ge \mu_{\tilde{S}}(s,x)$

then $\mu_{\tilde{A}}(x,y) = 1$. As a result, $\mu_{\tilde{A}}(x,y)$ can be defined as

$$\mu_{\tilde{A}}(x,y) = \frac{\int_{s}^{S_{\max}} \left[\mu_{[\tilde{S}]}(s,y) \wedge \mu_{\tilde{S}}(s,x) \right] \mathrm{d}s}{\int_{S_{\min}}^{S_{\max}} \mu_{\tilde{S}}(s,x) \mathrm{d}s} \tag{45}$$

where "∧" means "minimal". Thus, the determination of the membership function of the fuzzy event $\tilde{A}$ becomes the calculation of the membership functions of the fuzzy random generalized stress $\tilde{S}$ and fuzzy random generalized strength $\tilde{R}$ as in the form 2 model. Generally speaking, by using a system analysis method and applying the Extension Principle [30], the membership function $\mu_{\tilde{S}}(s,x)$ of $\tilde{S}$ can be calculated from the membership functions of the fuzzy load and the fuzzy geometric size of the structure. The membership function

Table 2. The membership function values of fuzzy safety state

S_{peak}	140	160	180	200	210	220	230	240
$\mu_{\tilde{A}'}(S_{\text{peak}})$	1	1	1	0.9995	0.9904	0.9401	0.8027	0.5397
S_{peak}	250	260	270	280	290	300	310	320
$\mu_{\tilde{A}'}(S_{\text{peak}})$	0.2699	0.1145	0.0367	0.0062	0.0004	0	0	0

$\mu_{\tilde{R}}(s,y)$ of $\tilde{R}$ could be in the form of semi-trapezoid distribution, or semi-normal distribution. By using the probability formula of a fuzzy event, the general expression for the reliability of a structure is

$$P_r = P(\tilde{A}) = \int_{R_{\min}}^{R_{\max}} \int_{S_{\min}}^{S_{\max}} \mu_{\tilde{A}}(x,y)p(x,y)\mathrm{d}x\mathrm{d}y \,, \qquad (46)$$

where $p(x,y)$ is the joint pdf. When only the randomness of generalized stress and fuzziness of allowable interval is taken into account, then

$$P_r = P(\tilde{A}) = \int_{S_{\min}}^{S_{\max}} \mu_{\tilde{A}}(x)p(x)\mathrm{d}x \qquad (47)$$

5.3. Simply supported beam under stress

Consider again the simply supported beam shown in Fig.3. Assume all random variables are normally distributed, that is

q~N(110,7^2) N/mm,
l~N(3600,150^2) mm,
b~N(120,10^2) mm,
h~N(240,10^2) mm.

Suppose the beam is made of #45 steel (a steel in China) and the membership function of its allowable bending stress is

$$\mu_{[\tilde{S}]}(s) = \begin{cases} 1 & , s \le 160 \\ (200-s)/(200-160), & 160 \le s \le 200 \\ 0 & , s > 200 \end{cases}$$

Now we determine the reliability of this simply supported beam. It is known from the strength of materials that the maximum stress of this simply supported beam is $S = 0.75ql^2/bh^2$. Because q, l, b and h are all normal random variables, the mean and standard deviation of the maximum stress are $\mu_S = 155\text{MPa}$, $\sigma_S = 24\text{MPa}$, respectively. It is known that the stress is a random variable and follows the normal distribution S~N(155,24^2) MPa. For the strength, only its fuzziness is taken into account. Therefore,

$$\mu_{\tilde{A}}(x) = \mu_{[\tilde{S}]}(s)|_{s=x} = \begin{cases} 1 & , \ x < 160 \\ \dfrac{200-x}{200-160} & , \ 60 \le x \le 200 \\ 0 & , \ x > 200 \end{cases}$$

where its domain is $\left[[\tilde{S}]_{\min}, [\tilde{S}]_{\max} \right] = [160, 200]$. The corresponding reliability is

$$P_r = P(\tilde{A}) = \int_{-\infty}^{+\infty} \mu_{\tilde{A}}(x)p(x)\mathrm{d}x$$

$$= \int_{S_{\min}}^{160} \frac{1}{\sqrt{2\pi}\times 24} e^{-\frac{(x-155)^2}{2\times 24^2}} \mathrm{d}x + \int_{160}^{200} \frac{200-x}{40} \frac{1}{\sqrt{2\pi}\times 24} e^{-\frac{(x-155)^2}{2\times 24^2}} \mathrm{d}x$$

$$= 0.825$$

6. Quantitative analysis of the influence of fuzzy factor on reliability

From the above mentioned analysis, we know that the reliability of the structure is conducted on the fuzzy safety state (fuzzy safety criteria). Moreover, the reason for introduction of the fuzzy safety state is that the structure state (functioning and failure) is ambiguous when the structure is on the limit state boundary ($z = 0$).

When the membership function of the fuzzy safety state follows the rising half-trapezoidal distribution, according to Eq. (15) with $z = 0$, we have

$$\mu_{\tilde{A}}(z)\big|_{z=0} = -\frac{a_1}{a_2 - a_1} \,. \qquad (48)$$

When the membership function of fuzzy safety state follows the rising half-ridge distribution, according to Eq. (24) with $Z = 0$, we have

$$\mu_{\tilde{A}}(z)\big|_{z=0} = \frac{1}{2} + \frac{1}{2}\sin\frac{\pi}{a_2 - a_1}\left(-\frac{a_1 + a_2}{2}\right). \qquad (48)$$

Let $\mu_{\tilde{A}}(z) = \alpha$ and $\alpha = [0,1]$, and then α is called the degree of confidence or confidence level of the structure safety when $z = 0$. The larger α becomes, the higher confidence level of structure safety at $z = 0$ will be. It is significant in practical applications to view this parameter as the criterion of confirming the fuzzy region, since the safety criterion is generally constructed on experiments or statistics. If statistics is quite comprehensive and experiment is highly reliable, then the safety criterion constructed on them has a higher confidence level. This shows that the fuzzy safety criterion can mostly consider uncertainty factors of human cognition. However, the conventional safety criterion is incapable of possessing this strong point. For a half-trapezoidal distribution, let $-\frac{a_1}{a_2 - a_1} = \alpha$ then

$$a_2 = \left(1 - \frac{1}{\alpha}\right)a_1 \qquad (49)$$

For a half-ridge distribution, let $\frac{1}{2} + \frac{1}{2}\sin\frac{\pi}{a_2 - a_1}\left(-\frac{a_1 + a_2}{2}\right) = \alpha$

then

$$a_2 = \frac{2\sin^{-1}(2\alpha - 1) - \pi}{2\sin^{-1}(2\alpha - 1) + \pi} a_1 \qquad (50)$$

That is to say, given a confidence level α, there exists the quantitative relation to a certain degree between the upper tolerance and lower tolerance, e.g., Eq. (49) and Eq. (50). Therefore, in case a_1 is given, a_2 can be confirmed accordingly. After parameter α is introduced, the variational relation of the fuzzy reliability of structure with the tolerance is as follows

$$P_r = f(\alpha, a_1) \qquad (51)$$

6.1. Simply supported beam under uniformly distributed load

A simply supported beam under uniformly distributed load is shown in Fig. 3. All basic random variables are assumed to follow normal distributions, i.e. $q \sim \text{N}(210, 7^2)$N/mm, $l \sim \text{N}(4000, 150^2)$mm, $b \sim \text{N}(120, 10^2)$mm, $h \sim \text{N}(240, 10^2)$mm. The beam is #45 steel and its strength is assumed to follow a normal distribution, i.e. $R \sim \text{N}(623, 23^2)$MPa. The membership function of the fuzzy safety state $\tilde{A}$ with proportional transition is

$$\mu_{\tilde{A}}(z) = \begin{cases} 0 & , \ z < a_1 \\ \dfrac{z - a_1}{a_2 - a_1}, & a_1 \le z \le a_2 \\ 1 & , \ z > a_2 \end{cases} \qquad (52)$$

with $a_2 = \left(1 - \frac{1}{\alpha}\right)a_1$. The relationship between the reliability and the tolerance is shown in Table 3, and the reliability curves are illustrated in Fig. 10. When the membership function of the fuzzy safety state $\tilde{A}$ is assumed to follow a half-ridge distribution, i.e.

$$\mu_{\tilde{A}}(z) = \begin{cases} 0 & , \ z < a_1 \\ \dfrac{1}{2} + \dfrac{1}{2}\sin\dfrac{\pi}{a_2 - a_1}\left(z - \dfrac{a_1 + a_2}{2}\right), & a_1 \le z \le a_2 \\ 1 & , \ z > a_2 \end{cases} \qquad (53)$$

where $a_2 = \dfrac{2\sin^{-1}(2\alpha - 1) - \pi}{2\sin^{-1}(2\alpha - 1) + \pi} a_1$. The relationship between the reliability and the tolerance is shown in Table 4, and the reliability curves are illustrated in Fig. 11.

We can arrive at conclusions by analyzing Table 3, Table 4, and Figs. 10-11,

1) The reliability of structures is associated with the confidence level α and tolerance a_1 besides basic random variables of structures. With the increasing of confidence level, the reliability of structures increases continuously. However, when confidence level is close to 1, the rate of increasing becomes low. This is consistent with the qualitative analysis.
2) When the confidence level α is lower, the increasing rate of reliability reduces along with the increasing tolerance. Moreover, when confidence level α is higher, the decreasing rate of reliability reduces along with the increasing of tolerance.
3) When the confidence level α is lower, the value of tolerance a_1 should not be too large, and when confidence level α is higher, the greater value of tolerance a_1 can be selected. This is also consistent with the qualitative analysis.

7. Conclusion

Fuzziness always exists in actual structure analysis. Since it is impossible to analyze the influence of fuzziness on structural reliability using the conventional reliability method, the evaluation of the structural reliability using the conventional reliability method cannot completely describe the reality. This work presents an investigation on reliability method of structures involved with fuzziness. The fuzzy safety state of a structure is defined by the state variable, fuzzy random

Table 3. The relation of reliability and tolerances

α \ P_r \ a_1	-0.05×623	-0.1×623	-0.15×623	-0.2×623	-0.25×623	-0.3×623	-0.35×623	-0.4×623	-0.45×623	-0.5×623
0.5	0.999990	0.999971	0.999820	0.999271	0.997346	0.992250	0.981452	0.962851	0.936768	0.906099
0.6	0.999997	0.999991	0.999972	0.999919	0.999766	0.999309	0.998553	0.996859	0.993800	0.988653
0.7	0.999998	0.999996	0.999992	0.999986	0.999973	0.999956	0.999892	0.999791	0.999622	0.999315
0.8	0.999998	0.999998	0.999998	0.999997	0.999996	0.999993	0.999991	0.999987	0.99998	0.999973
0.9	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999999	0.999998	0.999998	0.999998

Table 4. The relation of reliability and tolerances

α \ P_r \ a_1	-0.05×623	-0.1×623	-0.15×623	-0.2×623	-0.25×623	-0.3×623	-0.35×623	-0.4×623	-0.45×623	-0.5×623
0.5	0.999997	0.999983	0.999920	0.999677	0.999094	0.997647	0.994349	0.987739	0.978378	0.964035
0.6	0.999997	0.999991	0.999975	0.999929	0.999803	0.999506	0.998900	0.997781	0.995721	0.992396
0.7	0.999997	0.999996	0.999992	0.999981	0.999966	0.999911	0.999824	0.999674	0.999332	0.998745
0.8	0.999998	0.999998	0.999997	0.999996	0.999993	0.999988	0.999980	0.999961	0.999958	0.999909
0.9	0.999999	0.999999	0.999999	0.999999	0.999999	0.999998	0.999997	0.999996	0.999995	0.999994

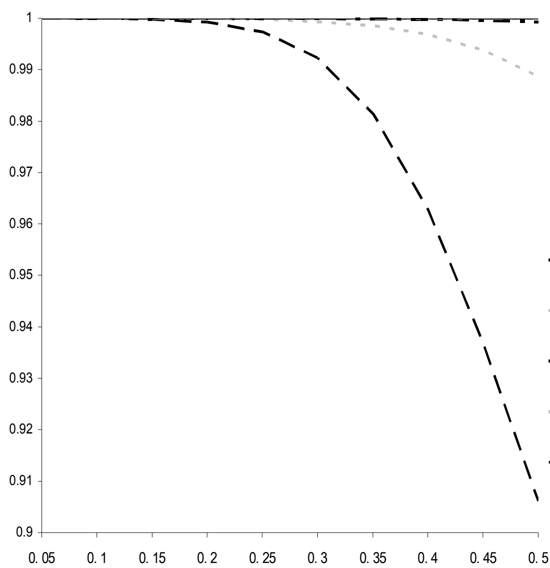


Fig. 10. Relationship of fuzzy reliability and tolerances

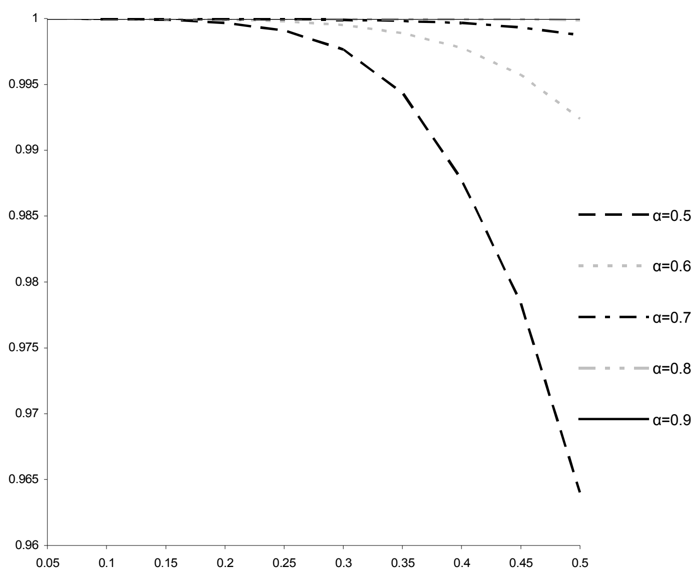


Fig. 11. Relationship of reliability and tolerances

allowable interval or fuzzy random generalized strength. Determination of the membership function of the fuzzy safety state is the key to the proposed method. This work concerns the fuzziness of the safety criterion, the fuzziness and randomness of generalized stresses and generalized strengths. The membership function of the fuzzy safety state is defined, and the structural reliability analysis method using fuzzy sets theory is proposed. Several examples are used to illustrate the proposed method.

Acknowledgement: *This research was partially supported by the National Natural Science Foundation of China under the contract number 51075061, and the National High Technology Research and Development Program of China (863 Program) under the contract number 2007AA04Z403.*

8. References

1. Blockley D I. The nature of structural design and safety. Chichester: Ellis horwood, 1980.
2. Blockley D I. Aprobabilistic paradox. Journal of Engineering Mechanics 1980; 106(EM6): 1430-1433.
3. Blockley D I. Risk based structural reliability methods in context. Structural Safety 1999; 21: 335-348.
4. Brown C B. A fuzzy safety measure. Journal of the Engineering Mechanics Division 1979; 105(5): 855–872.
5. Chou K C, Yuan J. Fuzzy-Bayesian approach to reliability of existing structures. J. Struct. Engrg. 1993; 119: 3276-3290.
6. Huang H Z, An Z W. A discrete stress-strength interference model with stress dependent strength. IEEE Transactions on Reliability 2009; 58(1): 118–122.
7. Huang H Z. Reliability analysis method in the presence of fuzziness attached to operating time. Microelectronics and Reliability 1995; 35(12): 1483–1487.
8. Huang H Z, Wang Z L, Li Y F, Huang B, Xiao N C, He L P. A non-probabilistic set model of structural reliability based on satisfaction degree of interval. Mechanika 2011; 17(1): 85–92.
9. Huang H Z, Tong X, Zuo M J. Posbist fault tree analysis of coherent systems. Reliability Engineering and System Safety 2004; 84(2): 141–148.
10. Huang H Z, Zuo M J, Sun Z Q. Bayesian reliability analysis for fuzzy lifetime data. Fuzzy Sets and Systems 2006; 157(12): 1674–1686.
11. Huang H Z, Li H B. Perturbation finite element method of structural analysis under fuzzy environments. Engineering Applications of Artificial Intelligence 2005; 18(1): 83–91.
12. Kapur K C, Lamberson L R. Reliability in Engineering Design. New York: John Wiley & Sons, 1977.
13. Knight Ridder/Tribune Business News. The nature of Colombia, Challenger shuttle disasters differ. The Dallas Morning News, April 9, 2003.
14. Liu Y, Qiao Z, Wang G Y. Fuzzy random reliability of structures based on fuzzy random variables. Fuzzy Sets and Systems 1997; 86: 345-355.
15. Moller B, Graf W, Beer M. Safety assessment of structures in view of fuzzy randomness. Computers and Structures 2003; 81: 1567-1582.
16. PBI Media, LLC. Pair of mechanical failures caused engine damage on E-4B. C4I NEWS, June 10, 2004.
17. Pugsley A G. The prediction of the proneness to structural accidents. Structural Engineer 1973; 51(3): 195-196.
18. Savchuk V P. Estimation of structures reliability for non-precise limit state models and vague data. Reliability Engineering and System Safety 1995; 47(1): 47–58.
19. Savoia M. Structural reliability analysis through fuzzy number approach, with application to stability. Computers and Structures 2002, 80: 1087–1102.
20. Schueller G I, Bucher C G, Baurgund U, Ouypornprasert W. On efficient computational schemes to calculate structural failure probabilities. Probabilistic Engineering Mechanics 1989; 4: 10–18.
21. Shiraishi N, Furuta H. Reliability analysis based on fuzzy probability. J. Engrg. Mech. 1983; 109: 1445–1459.
22. The McGraw-Hill Companies, Inc. CH-149 damage will have no impact on US101 bid. Aerospace Daily & Defense Report, October 21, 2004.
23. Tsukamoto Y, Tarano T. Failure diagnosis by using fuzzy logic. Proc. IEEE Conf. on Decision Making and Control, 1977: 1390–1395.
24. Wang Z L, Huang H Z, Du X. Optimal design accounting for reliability, maintenance, and warranty. Journal of Mechanical Design, Transactions of the ASME 2010; 132 (1): 011007-1-011007-8.
25. Wang Z L, Huang H Z, Liu Y. A unified framework for an integrated optimization model under uncertainty. Journal of Mechanical Design, Transactions of the ASME 2010; 132(5): 051008-1-051008-8.
26. Wang Z L, Li Y F, Huang H Z, Yu Liu. Reliability analysis of structure for fuzzy safety state. Intelligent Automation and Soft Computing 2012; 16(3): 215–242.
27. Wang G Y. Theory of soft design in engineering. Beijing: Science Press, 1992.
28. Wu W J, Huang H Z, Wang Z L, Li Y F, Pang Y. Reliability analysis of mechanical vibration component using fuzzy sets theory. Eksploatacja I Niezawodnosc - Maintenance and Reliability 2012, 14(2): 130–134.
29. Yao J T P. Damage assessment of existing structures. Journal of the Engineering Mechanics Division 1980; 106(4): 785-799.
30. Zadeh L A. Probability measures of fuzzy events. J. Math. Anal. App1. 1968; (10): 421–427.
31. Zadeh L A. The concept of linguistic variable and its application to approximate reasoning. Information Science 1975; 8: 199-249; 8: 301-357; 9: 43–80.
32. Zhang X L, Huang H Z, Wang Z L, Xiao N C, Li Y F. Uncertainty analysis method based on a combination of the maximum entropy principle and the point estimation method. Eksploatacja i Niezawodnosc - Maintenance and Reliability, 2012, 14(2): 114–119.

Prof. Hong-Zhong HUANG, 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:
JURECKI R, JAŚKIEWICZ M, GUZEK M, LOZIA Z, ZDANOWICZ P. Driver's reaction time under emergency braking a car – Research in a driving simulatory. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 295–301.

Rafał S. JURECKI
Marek JAŚKIEWICZ
Marek GUZEK
Zbigniew LOZIA
Piotr ZDANOWICZ

DRIVER'S REACTION TIME UNDER EMERGENCY BRAKING A CAR – RESEARCH IN A DRIVING SIMULATOR

CZAS REAKCJI KIEROWCY W WARUNKACH AWARYJNEGO HAMOWANIA SAMOCHODU – BADANIA W SYMULATORZE JAZDY SAMOCHODEM*

Extract. The paper presents results of research on the reaction time of drivers of motor vehicles in case of accident risk. These studies have been conducted in an autoPW driving simulator - research project N509 016 31/1251. During a simulated pre-accident situation, a test driver, as he was trying to avoid a collision with an obstacle, has been forced to perform a braking manoeuvre. 107 people of different age and driving experience have been studied. Various scenarios describing risk situations (the speed of a car driven by a tested driver, the distance between the vehicles at the initial moment) have been considered.Reaction times were determined at the controls of the vehicle and they have been presented in a time to collision function. The presented results provide important information on the analysed subject (for instance reaction time dependence on risk time). They have also confirmed the usefulness of this type of simulation studies in connection with the possibility of fairly accurate reproduction of the environment appearance and an event scenario in relation to possible real situations.

***Key words**: tests of drivers, drivers' reaction time, risk time, car driving simulators.*

W publikacji przedstawiono wyniki badań dotyczące czasu reakcji kierowców pojazdów samochodowych w sytuacjach zagrożenia wypadkowego. Badania te zostały wykonane w symulatorze jazdy samochodem autoPW w ramach projektu badawczego N509 016 31/1251. W zainscenizowanej sytuacji przedwypadkowej, badany kierowca próbując uniknąć zderzenia z przeszkodą, zmuszony był do wykonania manewru hamowania. Przebadano 107 osób różniących się wiekiem i doświadczeniem w prowadzeniu samochodu. Rozważono różne scenariusze opisujące sytuacje zagrożenia (prędkość samochodu, którym kierował badany, odległość pomiędzy pojazdami w chwili początkowej). Wyznaczono czasy reakcji na elementach sterowania pojazdem i przedstawiono je w funkcji czasu ryzyka (ang. time to collision). Zaprezentowane wyniki dostarczyły istotnych informacji na analizowany temat (np. zależność czasu reakcji od czasu ryzyka). Potwierdziły też przydatność prowadzenia tego rodzaju badań w symulatorach w związku z możliwością dosyć wiernego odwzorowania wyglądu otoczenia i scenariusza zdarzeń w stosunku do możliwych rzeczywistych sytuacji.

Słowa kluczowe: badania kierowców, czas reakcji kierowców, czas ryzyka, symulatory jazdy samochodem.

1. Introduction

According to statistical data, for instance [16, 19, 20], the most common cause of road accidents are persons driving the vehicles. It has been estimated that approximately 70% of accidents in Poland is a result of incorrect operation of a driver. Reaction time has been among the many features that characterize its function in the event of accident risk on a road. In short, it can be defined as the period from the time a threat appears to the moment a driver takes certain actions on vehicle controls to avoid an accident.

This parameter has a direct impact on the course of an accident situation, and road safety at the same time. On the other hand, it is one of the basic data adopted in the calculations carried out by experts from automotive and traffic engineering (and court experts) in the process of reconstructing accidents. The results of this analysis depend on values adopted in these calculations, which in turn, influence court's decisions concerning guilt of participants of the event or its absence. From both points of view, research on reaction time are thus fully justified.

The paper presents selected results of extensive research, conducted throughout 2007-2010, whose primary purpose was to gather particular reaction times. These studies have been conducted by three teams: the Kielce University of Technology (leader), the Technical University of Warsaw and Cracow University of Technology, under project no. N509 016 31/1251 "Development and update of the reaction times database of vehicle drivers."

2. Justification of the need to test reaction time

Data on drivers' reaction times are crucial in the manuals and training materials for court experts and automobile engineering and traffic experts. Values presented in many publications differ significantly. However, as mentioned earlier, values adopted in the analysis by a court expert may strongly influence final decisions on drivers's fault.

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

These differences often result from a variety of research methodologies (tool - test environment, test methodology, the multiplicity and "composition" of a focus group, the way of presenting results). From the perspective of usefulness for the accident reconstruction, it should be noted that no method developed so far could clearly be regarded as the best.

Generally, there are three methods (research environment) used for this type of research:

- stands for qualifying tests of drivers (used in the laboratories of transport psychology),
- experiments on roads or research/test tracks,
- research in driving simulators.

Evaluating reaction time in psycho-technical research is one of the sets of tests performed in the laboratories of psychology which are aimed at assessing general capacity of a tested driver within the scope of driving a car. These studies are characterized by a number of years of worked-out methodology and methods of assessing their results, for instance [22]. Reflection meters are used in the case of reaction time. Reaction time is estimated as the period from the onset of the light or sound stimulus to the moment of pressing a button on the desktop.

In the case of data from experiments on road or test tracks, these are often the results of reaction to so-called simple stimulus (a single tone or light), but the way drivers reacted is also simplified - it has to work on one of the vehicle controls (pedal brake, hand brake lever or steering wheel), [2, 5, 21, 31, 32]. The results of such research are often published as a recommendations to the experts, for instance [2, 33].

In real road situations (for instance, driving off in the column on a motorway, where the reaction takes place to the light "stop" before the car), a driver reacts to a complex stimulus. However, in the literature prior to 10-15 years, it is difficult to find data concerning reaction times, in which both the stimulus and the reaction of a driver are complex (as in real life accident situations). So far, studies carried out complex reactions to stimuli, but they were often highly simplified situations. For instance, in [5, 21], in studies of the reaction to a complex stimulus stimulator's lights stuck on the windscreen were used. In recent years, studies on the roads or tracks rely more often on the implementation of specified accident scenarios that were considered representative. The authors in previous papers [6, 13, 25, 26] presented this type of testing.

At this point the "observation" study in real traffic should also been mentioned. These studies often rely on analysis of records of a camera (for instance, monitoring) placed on the roads [23]. Results of evaluation (reaction time) here, however, strongly depend on observer's arbitrary assumptions as to the time of the initial emergency situation, and therefore the stimulus.

Development of simulation techniques, increase of computer and systems productivity to generate images allowed to use virtual research environment for testing drivers - driving simulators. Their use increases the independence of weather conditions, promotes the growth of reproducibility of results and test conditions [6, 17, 22]. It also allows for implementation of virtually any scenario without jeopardizing safety. Moreover, the experience of the authors indicate a strong correlation of research results in a simulator and a real car on a test track [6, 13].

To conclude this brief review of the literature, it can be stated that there is a need for research in which the reaction times will be determined not to a stimulus or the system of simple stimuli, but to a certain simulated risk accident situation. The number of available results in the literature for this type of testing is low and comprise a selection of special cases.

3. Tests characteristics

The project carried out research for three selected scenarios for risk accident situations. They have been conducted both on Kielce Track, as well as in an autoPW driving simulator. A common feature of the scenarios was the obstacle appearing suddenly and with further reduced visibility. The scenarios differed in the type of an obstacle (car, pedestrian, a set of tractor-trailer), the nature of traffic obstacles and other elements on the road, which influenced the complexity of the situation. Over 100 people have been tested in each scenario. As already mentioned, both track and simulator tests have been performed on tested drivers. Moreover, in both research environments, tests were made for different values of initial risk time conditions (ratio of the distance from the obstacle to the speed of a car being driven by test driver).

This paper presents the results for the third scenario, which was performed in an autoPW driving simulator. Test sample has been shown in Figure 1.

In an attempt mapped a situation in which a truck (a set of tractor-trailer) passed perpendicularly through the crossroads of two-lane carriageway in such a way that both lanes were blocked. The view of a transverse road has been severely limited. The road barriers were set on the left lane with typical markings for road works. The task of drivers during the research was an attempt to avoid a collision with an obstacle, but the way to do it has not been been imposed on them. The method of arranging the situation in practice, eliminated the option of choosing another defensive manoeuvre than the emergency braking (so-called emergency brake). Such an accident situation scenario was introduced as a result of the analysis of previous work by the authors [6, 7, 25, 26] and consultations with experts in the field of traffic accident reconstruction. The experts and court experts stressed the need to conduct studies in which the only reaction is hard braking, because regardless of the nature of the accident situation, it is definitely the most common defensive manoeuvre used by drivers.

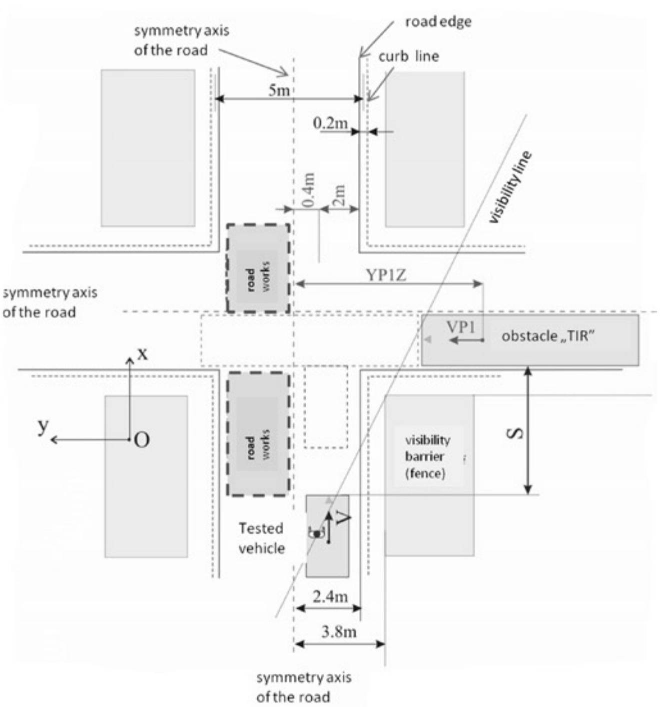


Fig. 1. Sample scheme for simulator tests (for scenario III)

For comparison, in the other scenarios that have been considered, the overall concept of the situation was similar – an obstacle which suddenly appears from the right side (car and pedestrian) at the crossroads. The avoidance manoeuvre was possible in the above-mentioned situations.

The study was carried out in an autoPW driving simulator, built and operated at the Department of Transport, Warsaw University of Technology [3, 17, 30]. It is the laboratory stand which allows the testing of drivers in staged traffic conditions, even in pre-accident traffic situations [3, 8, 9, 17, 30].

Basic elements of the simulator:

- natural drivers's cab coming from a middle class car with a set of equipment, the main screen and a secondary (side) on which is displayed (using a projector) an image seen through the windscreen (the viewing angle in the horizontal plane is over 90 degrees), the arrangement of position sensors of the vehicle controls (accelerator, brake, clutch, gear lever, switches, dashboard),
- computer chip simulator and data acquisition system for the exchange of information between the sensors and the computer system.

Vehicle motion is simulated in a computer system based on data from the sensors (which measure the values that characterise drivers), and data describing the vehicle and road conditions. The image seen by the driver is generated according to his actions and planned surrounding scenery, sound effects and state of the dashboard indicators. The activities of the tested driver are monitored and recorded. The autoPW simulator is a static type and the cab of the vehicle remains still during simulator operation (the driver does not feel body inertial stimuli). The mathematical model of the vehicle used in the simulator [17, 18] describes the dynamics of a vehicle. It has been positively verified experimentally for typical tests recommended by ISO [10]. Details of its construction can be found in sources [3, 17, 30].

Graphic capabilities of an auto PW driving simulator enable accurate graphical model a real crossroads. The crossroads of two streets in Warsaw has been selected for the above-mentioned studies. Images in the simulator were built based on photographic documentation of a crossroads and its surroundings. The geometric-spatial parameters and objects' colours have been reproduced in an exact way.

Studies have been conducted for the data corresponding to a FSO Polonez vehicle. Importantly, the simulator cab came from this car. The possibility of subsequent comparison of results obtained in the simulator for the study on Kielce Track where the same car was used, has also been of importance.

Tests have been conducted on 107 drivers:

- drivers under the age of 25 years - 76 people in a sample;
- drivers between the ages of 26 to 35 years - 11 people in a sample;
- drivers aged 36 to 45 years - 10 people in a sample;
- drivers aged over 46 years - 10 people in a sample.

J Risk time has been taken as the basic parameter characterizing the test. This parameter is described in earlier publications by the authors, for instance [6, 7, 13, 14, 25]. In some English-language publications can be found under the name of TTC (Time To Collision). The tests were conducted for 22 risk time values from 0.3 ÷ 3.6 obtained as combinations of vehicle speed and distance from the obstacle, that is:

- speed of the tested vehicle: 36, 40, 45, 50, 51.4, 60 and 65 km / h,
- distance from the vehicle when the driver noticed the obstacle: 5, 10, 20, 30, 40 and 50m.

Summary of test parameters has been shown in Table 1.

Figure 2 shows the example of research in the simulator (view "from above"). The sequence shown in these four images (frames) illustrating the course of a sample.

Tests for individual risk times (TTC times) have been randomly (from the standpoint of the test) mixed, while maintaining the same order for each driver. In this way (including the so-called "empty-pass") approximately 2500 tests have been conducted.

The values characterising the tested vehicle, obstacles and values describing drivers's actions on car controls have been recorded. Fig. 3 shows a record of selected parameters for the sample shown in fig. 2 (an interesting aspect was pulsating action of the driver on the brake pedal seen in fig. 3d). These recordings were the basis for the analysis of driver behaviour. This study has been limited to the size of main evaluated value - reaction time.

Fig. 2. Case study sample, risk of 1.8 s duration, vehicle speed 60 km/h, distance from an obstacle 30 m, the attempt ended without collision

4. The analysis of drivers' reaction times

Relationships characterized by reaction times were determined for the entire population of tested drivers. As described earlier, the specifics of the road situation in practice dictated one kind of behaviour: reducing traffic speed without the possibility of bypassing the obstacle. Therefore the following reaction times were determined:

- reaction time on the accelerator, understood as the time from the onset of the obstacle to the beginning of the leg off the ac-

Table 1 Realised risk time values for each sample during the test

Sample no.:	1s	2s	3s	4s	5	6s	7	8	9	10	11
Risk time, [s]	0,3	0,35	0,4	0,45	0,5	0,55	0,6	0,72	0,8	0,9	1,0
Speed of a tested vehicle V, [km/h]	60	51,4	45	40	36	65	60	50	45	40	36
Distance from an obstacle in the time of its appearance S, [m]	5	5	5	5	5	10	10	10	10	10	10
Sample no.:	12	13	14	15	16	17	18	19	20	21	22
Risk time, [s]	1,2	1,44	1,8	1,8	2,0	2,16	2,4	2,7	2,88	3,0	3,6
Speed of a tested vehicle V, [km/h]	60	50	40	60	36	50	60	40	50	60	50
Distance from an obstacle in the time of its appearance S, [m]	20	20	20	30	20	30	40	30	40	50	50

celerator pedal, in short: "acceleration" reaction time or mental reaction time ("trg" in Figure 4);
- psychomotor reaction time during braking, as the time from the emergence of obstacles to the onset of force on the brake pedal, in short: "brake" reaction time ("trh" in Figure 4);
- motor reaction time during braking, determined as the time from the start of removing the foot from the accelerator pedal to the onset of force on the brake pedal, in short: motor reaction time ("trm" in Figure 4).

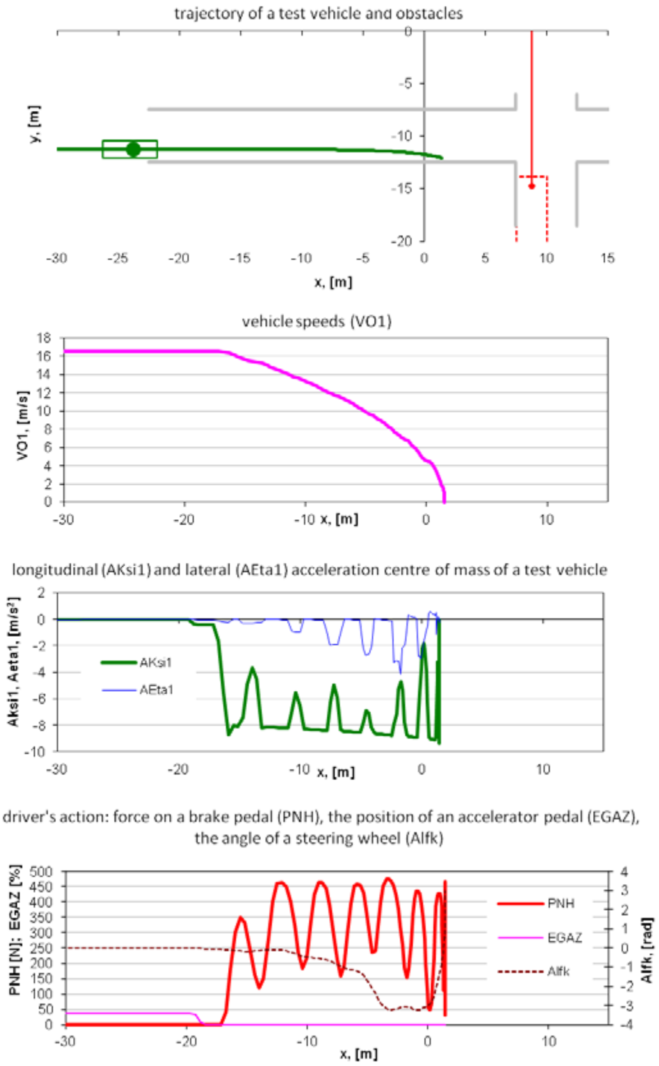


Fig. 3. Selected physical quantities recorded in a sample shown in fig. 2 (risk time 1.8 s), vehicle speed 60 km/h, distance from an obstacle 30 m vehicle speed 60 km/h.

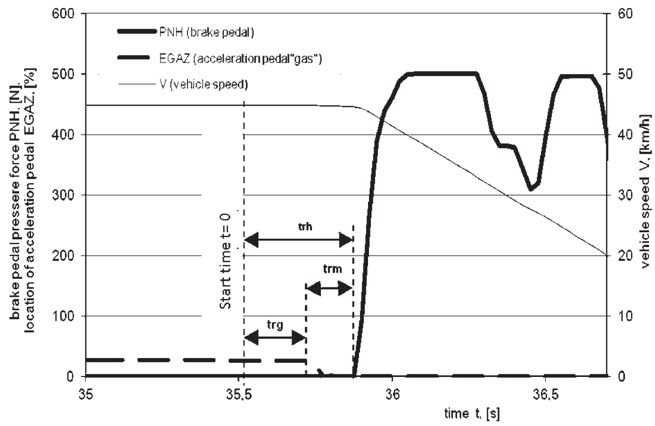


Fig. 4. Excerpts of selected waveforms recorded during a test - illustration of reaction times: trg – accelerator pedal reaction time "gas", trh - "brake" pedal reaction time , trm - motor reaction time

Figure 5 shows the "acceleration" pedal reaction times. The average values of this reaction time varied in the studied range of risk time from approximately 0.25 ÷. 0.6 s

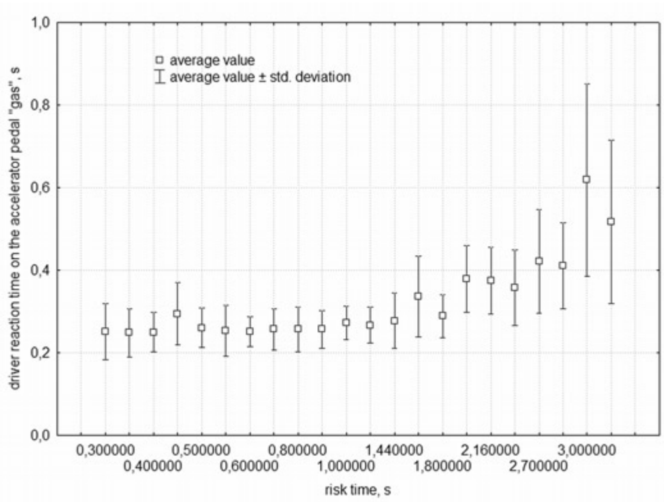


Fig.5. Acceleration pedal reaction time "gas" in a risk time function

In terms of the risk time value of 0.3 ÷ 1.2 s it can be considered that they have taken approximately constant level of approximately 0.25 s. In terms of risk time over 1.2 s, average reaction time values increased approximately linearly in a risk time function. Similar qualitative conclusions were formulated in relation to standard deviation. For risk time smaller than 1.2 s, the standard deviation was at an approximately constant level, amounting to 0.045 ÷ 0.06 s, and then

Table 2. Order of tests' realisation

Sample no. according to order of realisation:	1	2	3	4	5	6	7	8	9	10	11
(sample no. according to risk time – tab. 1):	(14)	(13)	(12)	(11)	(10)	(9)	(21)	(22)	(20)	(18)	(15)
Risk time [s]	1,8	1,44	1,2	1,0	0,9	0,8	3,0	3,6	2,88	2,4	1,8

Sample no. according to order of realisation:	12	13	14	15	16	17	18	19	20	21	22
(sample no. according to risk time – tab. 1):	(4)	(2)	(1)	(3)	(16)	(17)	(19)	(5)	(6)	(7)	(8)
Risk time [s]	0,45	0,35	0,3	0,4	2,0	2,16	2,7	0,5	0,55	0,6	0,72

began to rise to approximately 0.2 s, for the biggest analysed risk time of 3.6 s.

Deviations from described regularities occurred for risk times: 3.0 s 0.45s. For these tests the average and standard deviation values have taken a noticeably greater values than would appear from the trends described. Explanations of this phenomenon can be sought in the order of test implementation. As mentioned earlier, tests for subsequent risk time values have been mixed, but the same order for each of their drivers has been retained.

Schedule for each test are shown in table 2. This table identifies four specific tests, of whose distinctive feature was a big change of parameters compared to the previous test (mainly risk time).

The test of 3.0 s risk time was an attempt no. 7 and was performed after a few tests of increasingly less risk time value at the level of 1.0 to 0.8 sec. This could result in an additional surprise effect of a tested driver, as shown in the chart. A similar situation (large changes in the parameters of the sample) was at the test of 0.45 s risk time. This test was performed as no. 12 immediately after a series of tests with large values of the risk time: from 3.6 to 1.8 s

This time, the rapid changes in risk time also appeared with tests 2.0 s risk times (performed after several tests at risk times of 0.3-0.4 s) and 0.5 s (immediately after the tests of 2.7 s risk time) . Interestingly, in these cases there were no visible changes of average values and standard deviations of reaction time.

Interesting was also the observation of phenomena tentatively named by the authors' "first attempt to effect". Previously conducted studies at initial drive time values, which differed significantly from the described trend (were larger) [10, 11]. Here, this effect has not been observed. The reason for that could be that the participants already had previous experience with this type of testing. We could observe a factor associated with the "adaptation" to the simulator and the arrangement of the accident situation. Another factor that has been considered was a relative simplification of this scenario in relation to the scenarios I and II. In practice, the only option of a driver was a violent braking reaction.

Mental reaction time values were at a similar level to that achieved in previous scenarios (for instance [10, 11]), where the driver can also bypass the obstacle.

Figure 6 shows the average values and standard deviations of psycho-motor reaction time (reaction time on the brake pedal) in a risk time function. Qualitatively, dependences shown in this figure were similar to those shown in fig. 5, however, the values obtained were higher. Average values of psycho-motor reaction time varied in the range of 0.42 ÷ 0.92 s. Standard deviation values ranged from approximately 0.05 to 0.29 s. In this case the tendency to "stabilize" at a constant level of average values and standard deviations for the range of small risk time values was less evident. There was a large variation of concentration of reaction time distribution values around the average value. The measure of this concentration were the standard deviation. Its greatest value was almost 6-fold greater than the smallest value.

The described earlier deviation from the general trend also observed for the test of 3.0 s risk time and to a lesser extent also for .45 s test. Regarding the results of this study to the results obtained in other scenarios [10, 11] very similar waveforms were also found. In the case of scenarios I and III ranges changes the reaction time have almost overlapped.

Motor reaction time values have been shown in figure 7. Average values of this reaction varied, for the investigated time range of risk time, from 0.20 ÷ 0.34 s, while standard deviations ranged around 0.04 ÷ 0.15 s. For risk time greater than 1.0 s was observed a clear upward trend for both the average values, as well as for standard deviations. For smaller risk time values, both average values and standard deviations of the motor reaction time were approximately constant (average of around 0.2 s; deviation of 0.04 ÷ 0.06 s). An interesting fact was that in case of motor reaction, there was the lack of deviations from the general trend for the risk 3.0 times and 0.45 s.

Comparing the results for Scenario III with the effects of previous tests (scenarios I and II) waveforms were very similar both qualitatively and quantitatively [10, 11].

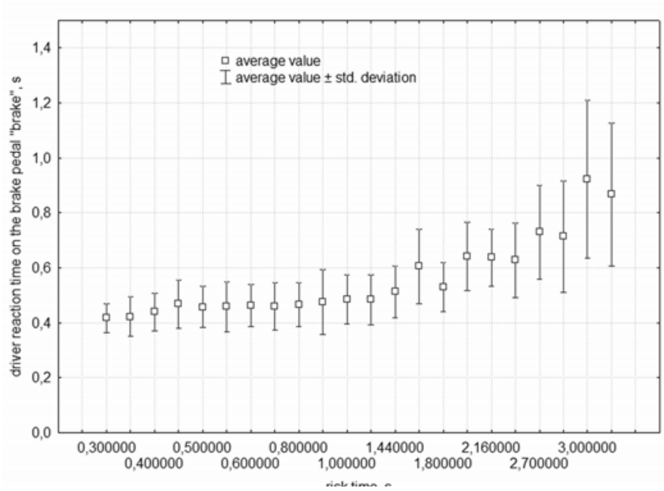


Fig. 6. Psychomotor reaction time (reaction time "brake") in a risk time function

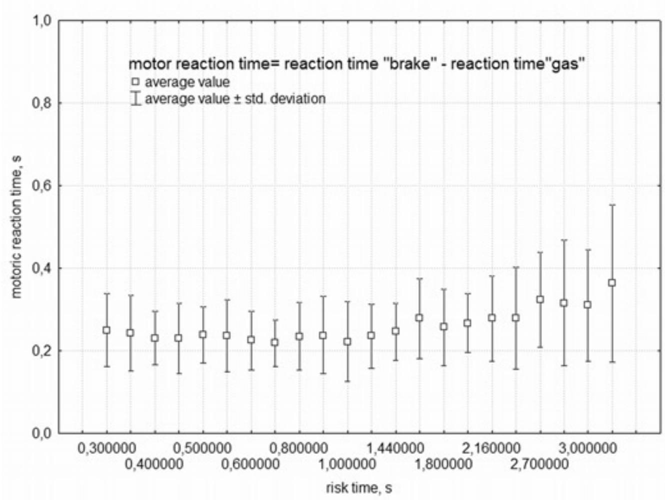


Fig. 7. The reaction times of drivers' motor reaction in risk time function

5. Summary

The presented results of drivers' tests for a particular scenario of an accident situation, support the view that the accident situation can be characterized by a risk time parameter. Similarly, as in previously conducted studies [6, 7, 8, 25] on smaller population of tested drivers and the research carried out for accident situations (the so called I and II scenarios [10, 11, 28]) the higher the risk time, the higher not only presented average reaction times but also the standard deviations that is their diversity increases - the dispersion of times distribution.

Another important observation is the quantitative similarity of the results obtained in the study of emergency braking to the results obtained in the scenarios (I and II). Average values of reaction times on the accelerator, the brake pedal and the motor reaction are at a similar level for particular risk times in all scenarios.

The last remark concerns the applied research tool - a driving simulator. The experience gained in the course of this study suggest that it might be a good tool to assess the behaviour of drivers in dangerous traffic situations [4]. Despite the many flaws (animated image, lack of inertial stimuli in an autoPW static simulator, "the artificiality" of the situation), the results obtained with its help provide important information about the behaviour of drivers in such situations. In this simulator you can perform tests that in the real world are not feasible or dangerous [1, 29]. On the other hand, the use of the simulator will be fully justified after checking whether there is a correlation between the results obtained on this unit and performed tests on a test track. Its presence has been demonstrated previously for the average reaction time [6, 13]. Analysis should be conducted to verify whether a correlation exists also for a number of tests carried out with one driver in both environments.

This paper concentrates on the drivers's reaction time associated with the implementation of the process of emergency braking. The amount of gathered data during the investigation allows a much wider analysis of driver behaviour. Such works have been in progress. They concern the variation of time on different parts of the vehicle controls (pedals: accelerator, clutch, steering wheel). The analysis also concern the way of reaction, its "intensity" and the efficiency for the parameters characterizing the event (scenario, risk time, speed, distance from an obstacle, etc.). Intra-individual characteristics of respondents - age, professional qualifications, experience in driving, etc. have also been taken into account. The partial results of these studies were published in [9, 10, 11, 15, 26, 27, 28].

Bibliography

1. Allen R. W, Cook M. L, Rosenthal T. J, Application of driving simulation to road safety. Advances in Transportation Studies. Publisher: University Roma Tre, ISSN: 1824-5463. Special Issue 2007, 12: 5-8.
2. Burckhardt M, Burg H. Die Brems-Reaktionsdauer von Pkw-Fahrer. Der Ferkehrsunfall 1981, 12: 224-235.
3. Chodnicki P, Guzek M, Lozia Z, Mackiewicz W, Stegienka I. Statyczny symulator jazdy samochodem autoPW, wersja 2003. Zeszyty Naukowe Politechniki Świętokrzyskiej. Mechanika 2004; 79: 157-164.
4. De Winter J. C.F, De Grott S, Mulder M, Wieringa P. A, Dankelman J, Mulder J.A, Relationships between driving simulator performance and driving test results. Ergonomics 2009, 52(2): 137-153.
5. Fröming R. Assesment of Integrated Pedestrian Protection System. PhD thesis. Fortschritt-Berichte VDI, Reihe 12, Verkehrstechnik/Fahrzeugtechnik 2008,681.
6. Guzek M, Jurecki R, Lozia Z, Stańczyk T. L. Comparative analyses of driver behaviour on the track and in virtual environment. Driving Simulation Conference Europe, DSC 2006 Europe, Paris, October 2006: 221-232.
7. Guzek M, Jurecki R, Lozia Z, Stańczyk T. L. Analysis of drivers on the track and in virtual environment. Zeszyty Naukowe Politechniki Świętokrzyskiej, Kielce 2006r., 84; 195-206.
8. Guzek M, Jurecki R, Lozia Z, Stańczyk T. Analysis of drivers' behaviour in pre-accident situations conducted in a driving simulator. Conference of the Forensic Study Institute IES, Szczyrk 27-29 października 2006r.: 157-171.
9. Guzek M, Jurecki R, Lozia Z, Stańczyk T. L, Zdanowicz P. Research on behaviour of drivers in accident situation conducted in car driver simulator" Journal of KONES. Powertrain and transport 2009; 16: 173-184.
10. Guzek M, Jurecki R, Lozia Z, Karendał M, Zdanowicz P. Badania reakcji kierowców na pieszego wychodzącego z prawej strony, realizowane w symulatorze jazdy samochodem; Autobusy. Technika, Eksploatacja, Systemy Transportowe 2010; 6, wersja CD.
11. Guzek M, Jurecki R, Lozia Z, Stańczyk T. L, Zdanowicz P. Analysis of drivers' reactions to a vehicle approaching from the right side, carried out in a driving simulator, Warsaw University of Technology Scientific Papers 2010; 77, 129-139.
12. Guzek M, Lozia Z, Pieniążek W. Experimental verification of a simulation model used in a driving simulator. Warsaw University of Technology Scientific Papers 1999; 34, 69-87.
13. Jurecki R. S. Modeling driver's behaviour in preaccident situations. The doctoral dissertation. Kielce University of Technology, Faculty of Mechatronics and Machine Building, Kielce 2005r.
14. Jurecki R, Stańczyk T. L. Driver model for the analysis of pre-accident situations. Vehicle System Dynamics 2009; 47 (5), 589-612.
15. Jurecki R, Stańczyk T. L. The test methods and the reaction time of drivers. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2011; 3: 84-91.
16. Kisilowski J, Zalewski J. Selected problems of analysis of the causes of road accidents in Poland in the years 1995-2004. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2008; 1: 37-43.
17. Lozia Z. Driving simulators. WKŁ, Warszawa, 2008.
18. Lozia Z. Simulation model of movement and dynamics of a two-axle vehicle, used in the simulator. Papers of the Institute of Vehicles - Warsaw University of Technology, Warszawa 1999; 4(34): 37-51.
19. Information materials of the Police Main Headquarters, Road accidents – annual reports, [online: http://dlakierowcow.policja.pl/, access: 2011-11-25].
20. Information materials of the National Board of Traffic Safety (KRBRD) [online: http://www.krbrd.gov.pl/stan_brd/, access: 2011-11-25].
21. MA T, Krulec R, Batista M, Bogdanović L. The driver reaction time measurement experiences. Innovative Automotive Technology – IAT'05, Bled,21st-22nd April 2005.
22. McGehee D.V, Mazzae E.N, Baldwin G.H.S. Driver reaction time in crash avoidance research: validation of a driving simulator study on a test track. Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society (IEA 2000), San Diego/USA 2000: 6.
23. Muttart J.W. Driver reaction in various environments estimated empirically. Materials of the conference entitled "Problems of traffic accident reconstruction." Publisher: the Forensic Institute Kraków, 2004.
24. Rotter T. Methodology of psychological tests of drivers - a revised version, the Institute of Road Transport, Warszawa 2003r.
25. Stańczyk T. L, Jurecki R. The reaction times of drivers in emergency accident situations. Materials of III Conference entitled "Development of automotive technology vs. communication insurance" Radom, 2006. Wyd. WSB im. J. Chrapka: 321-348.
26. Stańczyk T. L, Jurecki R. Fahrereaktionszeiten in Unfallrisikosituationen – neue Fahrbahn- und Fahrsimulatorversuche, Verkehrsunfall und Fahrzeugtechnik 2008; 07-08: 235 – 246.
27. Stańczyk T, Jurecki R, Pieniążek W, Jaśkiewicz M, Karendał M, Wolak S. Analysis of drivers' reactions to a vehicle approaching from the

right side, carried out on a car track. Papers of the Institute of Vehicles Warszawa 2010; 77: 307-320.
28. Stańczyk T. L, Lozia Z, Pieniążek W, Jurecki R. S. Analysis of drivers' reactions in simulated accident situations, Scientific Papers of Warsaw University of Technology - Vehicle Department, Warszawa 2010; 77: 27-52.
29. Strayer D. L., Drews F. A., Crouch D. J., A comparison of the cell phone driver and the drunk driver. Human Factors 2006; 48: 381. (http://hfs.sagepub.com).
30. autoPW driving simulator [online: http://www.it.pw.edu.pl/autopw/; access: 2011.11.30].
31. Törnros J. Effect of driving speed on reaction time during motorway driving. Accident Analysis and Prevention 1995, 27(4): 435-442.
32. Wang B, Abe M, Kano Y, Influence of driver's reaction time and gain on driver-vehicle system performance with rear wheel steering control systems: part of a study on vehicle control suitable for the aged driver, JSAE Review 2002; 23 (1): 75-82.
33. Road accidents. Court Expert Handbook. Collective paper. IES Kraków Publishing House 2002.

Rafał S. JURECKI, Ph.D., Eng.
Marek JAŚKIEWICZ, Ph.D., Eng.
Kielce University of Technology
Faculty of Mechatronics and Machine Building
Department of Motor Vehicles and Transportation
Aleja Tysiąclecia Państwa Polskiego 7, 25-143 Kielce, Poland
E-mail: m.jaskiewicz@tu.kielce.pl; rjurecki@tu.kielce.pl

Marek GUZEK, Ph.D., Eng.
Prof. Zbigniew LOZIA, Ph.D., Eng.
Piotr ZDANOWICZ, M.Sc., Eng.
Warsaw University of Technology
Faculty of Transport
Ul. Koszykowa 75, 00-662 Warszawa, Poland
E-mail: mgu@it.pw.edu.pl; lozia@it.pw.edu.pl; zdanowicz@it.pw.edu.pl

Article citation info:
ZHANG Z, GAO W, ZHOU Y, ZHANG Z. Reliability modeling and maintenance optimization of the diesel system in locomotives. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 302–311.

Zhisheng ZHANG
Wenke GAO
Yifan ZHOU
Zhiqiang ZHANG

RELIABILITY MODELING AND MAINTENANCE OPTIMIZATION OF THE DIESEL SYSTEM IN LOCOMOTIVES

MODELOWANIE NIEZAWODNOŚCI I OPTYMALIZACJA UTRZYMANIA RUCHU UKŁADU SAMOCZYNNEGO ZAPŁONU W LOKOMOTYWACH

Engine system is a prone-fault part in diesel locomotive and its malfunctions always occur regularly in different seasons in practice. However, the current maintenance policy in China has not attached deserving importance to seasonal influence, which is considered as one of the main causes for over/under-maintenance. To assess the current maintenance, in this study a double-fold Weibull competing risk model for summer and winter is developed using the real failure data (2008-2011) of locomotives from Urumqi Railway Bureau. Meanwhile, a new approach, termed as Approximately Combined Parameter Method (ACPM), is proposed to combine the initially estimated parameters into different folds, which can avoid a subjective determination of the model's parameters fold. After that, the combined parameters are used as initial values for maximum likelihood estimate (MLE) to achieve an accurate model. Necessary optimizations are introduced based on the chosen models. Results show that the maintenance period differs a lot between winter and summer, and the optimized maintenance can increase the availability and decrease cost more than the existing policy

Keywords: *diesel engine of locomotive, multiple Weibull competing risk model, maintenance optimization, approximately combined parameter method(ACPM).*

Układ silnikowy stanowi podatną na uszkodzenia część lokomotywy spalinowej, a w praktyce jego awarie występują zawsze regularnie w zależności od pory roku. Pomimo tego, obecna polityka obsługowa w Chinach nie przywiązuje wystarczającej wagi do wpływu pór roku, co uważa się za główną przyczynę nadmiernych lub niedostatecznych działań obsługowych. Aby ocenić bieżące działania obsługowe, w niniejszym artykule opracowano model zagrożeń konkurujących dla lata i zimy, oparty na połączeniu dwóch rozkładów Weibulla, wykorzystujący rzeczywiste dane o uszkodzeniach (2009–2011) lokomotyw używanych przez Agencję Kolejową Urumqui. Jednocześnie zaproponowano nowe podejście, o nazwie Approximately Combined Parameter Method (Metoda Przybliżonego Łączenia Parametrów, ACPM), które polega na łączeniu wstępnie obliczonych parametrów w różne wielokrotności, co pozwala na uniknięcie subiektywnego wyznaczania liczby parametrów modelu. W celu otrzymania dokładnego modelu, połączone parametry wykorzystuje się jako wstępne wartości w estymacji metodą największej wiarygodności. Konieczne optymalizacje wprowadza się na podstawie wybranych modeli. Wyniki pokazują, że letni okres obsługowy różni się zasadniczo od zimowego, a zoptymalizowana obsługa może zwiększyć gotowość systemu i zmniejszyć koszty utrzymania ruchu w większym stopniu niż dotychczasowa polityka obsługowa.

Słowa kluczowe: *lokomotywowy silnik Diesla, model konkurujących zagrożeń oparty na wielokrotnym rozkładzie Weibulla, optymalizacja utrzymania ruchu, metoda przybliżonego łączenia parametrów (ACPM).*

1. Introduction

The diesel engine system of locomotives must satisfy stringent reliability and availability requirements. Statistics shows that the system accounts for about 60% malfunctions and over 60% maintenance costs of diesel locomotives in China. Investigations indicate that some of them are caused by over/under-maintenance. The current maintenance policy of keeping periodical and condition-based maintenance under scheduled preventive maintenance is regarded as one of the major causes for over/under-maintenance. Therefore, optimizing the current maintenance policy is a must for enhancing the reliability and availability of diesel locomotives and their components.

So far, various approaches have been developed to improve the maintenance strategy for diesel locomotives, apart from improving the function of components. Some of them focus on condition-based maintenance, in these strategies, the maintenance duration or some aided decision are made through collecting actual technical state of key components based on monitoring [1, 3] or detecting information [1, 10]. Though potential failures of certain key components may be detected in time by this mode, it has not gained wide application to most of Chinese diesel locomotives as only a small number of components can be checked. Only few of them determine maintenance according to some parameters of key components. Lingaitis [8] et al. propose a method to determine the maintenance data using the state of fuel consumption of diesel locomotives, while the fuel consumption of diesel locomotive is easily influenced by some unpredictable rand factors, such as state of railway and traction weight and outside condition and so forth. Wei Di [14] employed a physical model by calculating the accumulative damage degrees of main generator according to plenty of operation records, and then determine their major maintenance period, and yet the physical model is restricted more because of the complexity of its failure mechanism. Little exists in the literatures to optimize common maintenance period of diesel locomotives based on the real failure data. In additional, The influence of environmental condition on the reliability of electromechanical equipments can not be ignored [11, 15]. Experience shows that malfunctions of the diesel engine system in locomotives always occur irregularly in different seasons in practice, which obviously influence the reliability of diesel engine system. While the current maintenance policy in China does not attach deserving importance to the seasonal influence, and there is no research to investigate the severity of seasonal influence on the diesel engine system. Therefore, it is necessary to develop a model considering the seasonal condition to assess and optimize the maintenance for diesel locomotives.

Weibull Model is widely used in reliability modeling of electromechanical equipments and in maintenance optimization [2, 4, 12]. As a typical electromechanical equipment, diesel engine system is suitable for Weibull model. Dan Ling et al. [7] proposed a method using Nonlinear Least Squares (NLS) theory and quasi-Newton method to estimate the parameters of mixture Weibull model. Chanseok Park [13] used EM method to estimate parameters of incomplete data in Competing Risk model. Furthermore, graphic approach and MLE are commonly used to estimate parameters because numerical solution can not yield a close form solution in general [9]. R.J proposed a method to estimate initial parameters of Double [6] and n-fold [5] Weibull Competing Risk Model in reference [5, 6]. D. Bocchetti et al. [2] developed a reliability model for cylinder liners of marine diesel engines using the method. However, there are two points that should be noted in modeling of Weibull Competing Risk Model: a) the fold number, which may turn out to be multiple, should not be determined beforehand by observation; b) the termination of the algorithm is subjective when the author estimates the parameters of multifold Competing Risk Model. While there is no literature to further study or to develop a beneficial algorithm for computer programming.

In view of the above, as an example of DF4B diesel locomotive, two main parts are introduced in this paper. One is the process of modeling. In this part, two double-fold Weibull Competing Risk models for winter and summer are respectively developed using the real failure data of Korla Locomotive Sect of Urumqi Railway Bureau, to estimate the reliability and optimization of diesel engine maintenance. The influence of seasonal variables on reliability is firstly considered in modeling, and a new method named ACPM is proposed, which can combine the initially estimated parameters into different folds, prevent the fold of parameters from being subjectively determined, and facilitate computer programming, at the same time, which can offer several models for making a choice. To select the best model among all that we obtain, the Bayesian Information Criterion (BIC) evaluation is employed as a criterion in this paper. For the selected models, the running mileage of locomotives is divided into three phrases serving as references for optimization. Another is the process of the maintenance optimization. In this part, the maintenance optimization based on availability and cost respectively as well as on both is explored. Results show that the maintenance period differs a lot between winter and summer. In additional, the effect of preventive maintenance (PM) cost and minimal repair cost on maintenance period in cost-oriented optimization is discussed in detail.

The rest of the paper is organized as follows: Section 2 introduces the process of building a reliability model of a diesel engine system. Maintenance optimization and discussion of the results come in section 3. A brief summary is given in the last section.

Symbols:

$F(t)$	Function of system failure rate
$f(t)$	Function of system failure rate density
$R_w(t)$	Function of reliability in winter
$R_s(t)$	Function of reliability in summer
$r_w(t)$	Function of failure rate in winter
$r_s(t)$	Function of failure rate in summer
C_p	PM action cost
C_f	Minimal repair cost
T_p	Time for PM action
T_f	Time for minimal repair

2. Reliability Model of Diesel Engine System

2.1. Failure Analysis of Diesel Engine System

China has many series diesel locomotives with a total of 11041, DF4B that is produced in 1980s or earlier is one of them, and that is about 4300 amount for about 40% in the whole diesel locomotives in China. Because of the large demand of locomotives in China, DF4B diesel locomotives may still serve on main or branch railway line for a long period. In the daily operation, the maintenance cost and malfunctions become outstanding problems bothered railway enterprises. Therefore, in this paper, DF4B diesel locomotives are taken as an example to investigate their reliability and maintenance policy.

For diesel locomotives, fewer failures occur in the body of engine, cylinders, pistons, crankshafts or transmission gears in a maintenance interval of diesel engine system, and most of them are caused by fuel and lubricating oil system, cooling system, supercharger, inletting or exhausting system in daily operation. Statistics of DF4B diesel engine system shows that malfunctions of the fuel and lubricating oil system take up about 43% in the whole account. Among the fault-prone components are fuel pumps and pipelines, fuel injection pumps, injectors, combined regulators, fuel supply gears, fuel supply levers, oil pumps, etc. Failures resulting from cooling systems composed of high/low-temperature cooling devices account for 32% or so. Another 14% is taken by those from turbochargers, inletting and exhausting systems, key components that demand costly maintenance and are prone to damage like burnt bearings, broken blades and over-large inletting gaps. Common malfunctions are shown as turbocharger bearings burned, blades damaged and over large on inletting gaps. Other uncommon malfunctions account for around 11%, most of which are caused by mechanical wear or unexpected break.

2.2. Data Preprocessing

To reveal seasonal influences on the diesel engine system, the failure data are divided into the groups of winter and summer based on the local climates as seen in Tables 1 and 2, which include the annual faults statistics (2008-2011) of DF4B diesel locomotives from Urumqi Railway Bureau, and which are complete failure data, covering repair time, running mileage after repair, failed components and reasons.

Assume that the failure data follow the Weibull distribution, and their initial reliability can be attained from the Median Rank Estimates as Eq. (1). Then switch each ti into versus initial reliability R(ti) utilizing the Weibull Conversion as Eq. (2) and plot a Weibull Probability Plot (WPP), as shown in Fig. 1, from which it can be seen that the distribution curve are befitting with the Weibull competing risk model.

Table 1. Failure Data in Summer

Failure data	PM interval	Failure data	PM interval	Failure data	PM interval	Failure data	PM interval	Failure data	PM interval
8276	31689	5593	35875	12543	44676	22399	34420	29793	
5667	37973	5996	39418	13063	39338	19008	39462	35403	39875
27933	42973	6218	36954	13748	32261	24323		7617	
4032	44921	15600		7458	39543	17434	37747	30991	
21420		7015	30750	14244		24490		21162	45699
2904	31852	14892		6966	37598	24798	40858	4930	
21162		8168	38476	16592		35403		31956	42359
3318	46006	8426	40352	7919	38974	25412	40233	32006	38403
22985		24566		18877		25526	39552	31915	
3355	36420	8647	38403	236	40373	26500	35875	32234	37419
3738	38577	10207	42489	8559		3209	40373	32844	38465
4375	46286	23219		25793		27923		29270	
30990		10907	35858	19840	40142	19056		34005	42013
212	31480	11450	41868	20751	46467	28025	39875	26031	
4877		11947	39315	34363		34865		34676	42675
5442	36116	12022		10528	35465	28177	37849	19528	46265
30025		21207	46418	21891		1478	38250	1867	
4797	35190	32046		28673	36964	28704		28633	31168
4431	34461	35556	39705	19696	36220	39909	46467	42627	46474
29997		36635	40586	4799		40687	46555	37149	
460	37722	400	40286	32937	34670	375	36714		
28832		23425							

Table 2. Failure Data in Winter

Failure data	PM interval	Failure data	PM interval	Failure data	PM interval	Failure data	PM interval	Failure data	PM interval
2471		38001	38517	8029	39620	6996	37557	7392	41955
8865	38780	31355	38398	8865	41383	24414		37410	
22436		7771	40463	36981		26700	36560	26460	41024
41012	42522	1496		10976	41189	27831	37367	4711	45413
8956	38054	15447	44282	11137	31026	28665	40373	44848	
37042	41735	32582	40572	4330	37828	28811	41733	3024	41755
9961	38324	26861	40853	11743		41746	42546	7015	
32582		31658	39272	16335	40466	31129	36164	3601	39711
220	44274	29452	35332	18622	38154	31398	39418	3879	38452
21657	36276	28835	38951	19300	39418	31472	42675	11138	
33400	38861	6680	40601	19683	41489	32358	39902	5023	36560
2829	33794	3056	37577	22074	40150	32784	40373	2224	36714
29290	38517	375	40606	22985	37722	9211	37330	7051	41295
30284		6514		33817	36165	34199	37537	23681	
33970	36815								

Median Rank Estimates of Initial Reliability $R(t_i) = 1 - \frac{i - 0.3}{n + 0.4}$ (1)

Weibull Conversion $\begin{cases} x = \ln t_i \\ y = \ln(-\ln R(t_i)) \end{cases}$ (2)

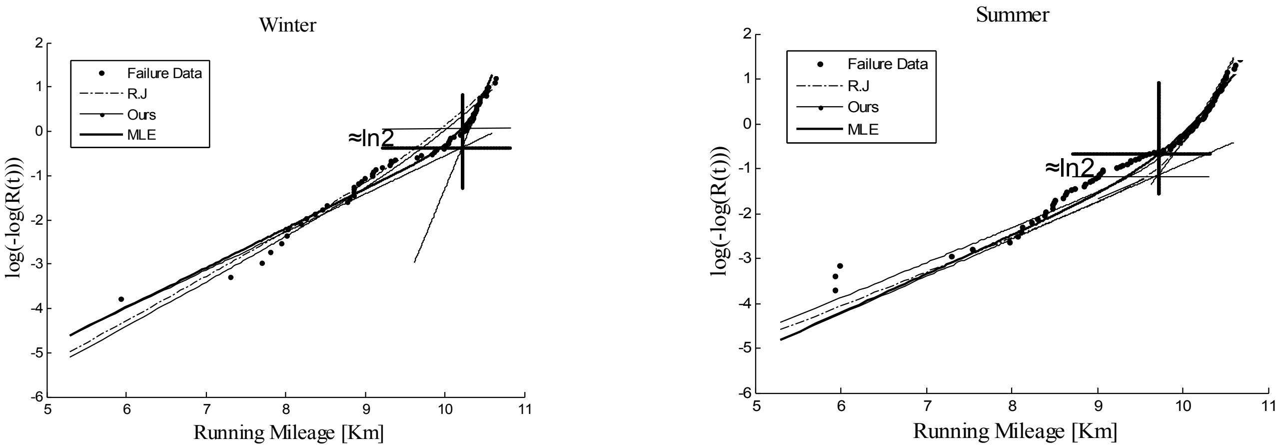


Fig. 1 WPP

2.3. Initial Parameter Estimation

Parameter estimation is important but usually difficult as methods like the maximum likelihood estimation cannot yield a close form solution in general. The numerical calculation and iteration method are needed. There are different methods which can be applied to model parameter estimation. Among them, the graphical approach such as the WPP method and MLE are of the common use. Jiang et al. have separately introduced methods to estimate parameters in Dual and Multiple Weibull competing risk model in literatures [5, 6]. The method can apply to large sample data, and can be regarded as an efficacious separation method often applying in engineering.

However, the terminal condition of the sub-sample separate algorithms that R.J proposed is judged personally, which may cause uncertainty of folds of estimated parameters, ensuring no optimal model and complicating computer programming. In this study, we set the termination using two approximately parallel lines, viz., the left asymptote of residual-data-fitted curve and the whole residual-data-fitted line. Let ka be the left asymptote slope of residual-data fitted curve, and kl be the slope of the residual-data fitted line, and let the algorithms end when $k_l \approx k_a$. Though it is still an approximate condition, among the estimated parameters several group βs are close when residual data is close to a straight line, and the distance among βs in the next section.

Applying the proposed terminal condition to separate sub-samples under the Matlab 7.0 step by step, the initial model of the diesel engine system is 3-fold and 5-fold for winter and summer respectively, as shown in Tables 3 and 4.

2.4. Approximately Combined Parameters

Further study is needed for accurate estimation as the initial models have been obtained by approximate means. As mentioned in the previous section, k_l is more close to k_a when the residual data distribute in a nearly straight line. In this study, the Hierarchical Clustering Method [16] is employed to approximately combine the initially separated parameters into different folds, from which a valid model is detected. This method is termed as approximately combined parameters (ACPM), the detailed procedure is illustrated in Fig. 2.

Let β_1 、⋯、β_n , denote the shape parameter, and η_1 、⋯、η_n denote the scale parameter, which are determined by sub-sample separation in the initial estimating process. Let $d_i = \beta_{i+1} - \beta_i$, and D_k is a distance matrix composed of d_i. The whole data is drafted to be divided into N ($N \le n$) categories, the sorting procedure is shown in Figure 2. Where the threshold value δ_0 is determined by the allowable error ς, and when $\min(t_i)^{\beta_i - \bar{\beta}} \le \varsigma$, assume approximately that $\beta_i \approx \bar{\beta}$,

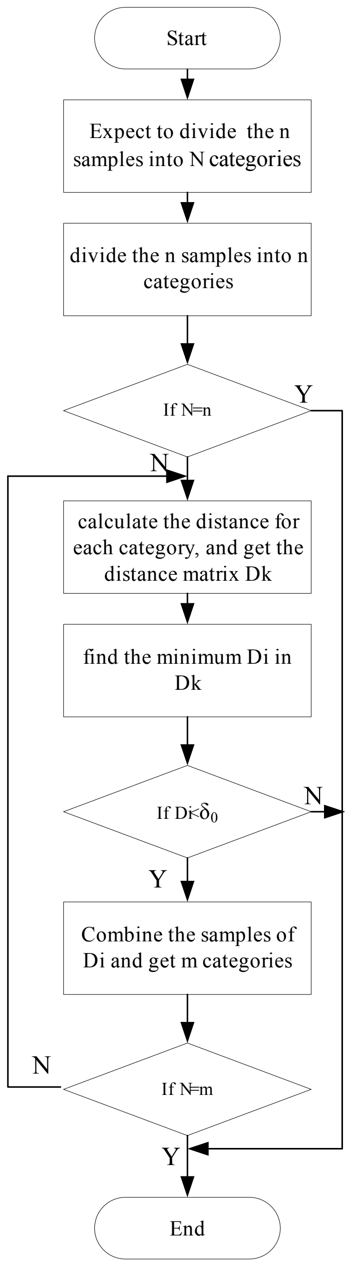


Fig. 2. Approximate Combine of Parameters

Table 3. Model Parameters of Diesel Engine System for Winter

β	η	BIC	Fold	Remark
0.9822 0.9931 2.2444	89980 56328 34289	1535.4	3	Initially estimated parameters
0.8794 1.0446 2.5826	92979 56327 34289	1462.9	3	Initially estimated parameters used as initial values for MLE
0.98765 2.2444	34732 34289	1526.9	2	ACPM
0.8745 5.9318	35199 34289	1356.3	2	Initially estimated parameters used as initial values for MLE
1.4065	1542	1523.8	1	See P.S.
1.1005	22911	1421.4	1	One-fold parameters used as initial values for MLE

Table 4. Model Parameters of Diesel Engine System for Summer

β	η	BIC	Fold	Remark
0.69483 0.76736 0.79204 0.86951 4.347	346010 140120 320010 337960 26499	1831.3	5	Initially estimated parameters
0.69611 0.82003 0.9791 1.1033 3.0132	140430 333070 319510 338230 26519	1849.4	5	Initially estimated parameters used as initial values for MLE
0.69483 0.7797 0.86951 4.347	346010 76278 337960 26499	1821.7	4	ACPM
0.95018 0.71451 1.238 3.0268	333070 76363 338230 26519	1814.9	4	Initially estimated parameters used as initial values for MLE
0.75141 0.86951 4.347	49296 337960 26499	1811.1	3	ACPM
0.7655 7.8098 4.9915	49175 338232 31075	1857.4	3	Initially estimated parameters used as initial values for MLE
0.78093 4.347	30692 26499	1799.9	2	ACPM
0. 86 3.1032	30239 26519	1794.1	2	Initially estimated parameters used as initial values for MLE
1.4941	219.7	1831.3	1	See P.S.
1.0529	22982	1807.1	1	One-fold parameters used as initial values for MLE

P.S. When parameters are combined as one-fold, the error is too large to be acceptable, which does not meet the requirement of ACPM algorithm, for the sake, we apply parameters that fit the original failure data in a straight line are chosen.

calculate η by Eq. (3) after the combination. Let $\beta_i - \bar{\beta} = \delta_0$ in the program.

From $$\sum(\frac{t}{\eta_i})^{\beta_i} = \sum\frac{t^{\bar{\beta}_j} t^{\beta_i-\bar{\beta}_j}}{\eta_i^{\beta_i}} \approx t^{\bar{\beta}_j}\sum\frac{1}{\eta_i^{\beta_i}}$$

We have $$\bar{\eta}_j = (\sum\eta_i^{-\beta_i})^{-\bar{\beta}_j^{-1}} \tag{3}$$

and $$\sum(\frac{t}{\eta_i})^{\beta_i} \approx (\frac{t}{\eta_j})^{\bar{\beta}_j} \tag{4}$$

The ACPM combine the estimated parameters into expected categories are indicated in Tables 3 and 4. Both the initially estimated and the approximately combined parameters are used as initial values for MLE to obtain the accurately estimated parameters, which are given in the two tables as well. The choosing of reasonable models is introduced in the next section.

2.5. BIC Evaluation of Models

Based on the ACPM and MLE, BIC is usually utilized to evaluate all the models for bigger sample size. The BIC evaluation of each model is calculated by Eq. (5), where N is the number of failure data, L is the Maximum Likelihood Function Value of Estimation Model, and k is the number of Parameters. All the BIC evaluation values are indicated in Tables 3 and 4, among them the smallest one is considered as the desirable model marked in different color.

$$BIC(k) = -2\ln L + k\ln N \tag{5}$$

As indicated by Tables 3 and 4 that ACPM is better than initially estimated model according to the BIC value. Fig. 1 shows the initially estimated model and the double-fold model for MLE and ACPM. According to the selected model that the intersection point (x_I,y_I) between the left asymptote and the right asymptote meet $R(x_I)-y_I\approx\ln 2$. Therefore, the selected model for winter and summer is denoted as function (6) and (7) respectively.

$$R_w(t) = \exp(-(t/35199)^{0.8754} - (t/34289)^{5.9318}) \tag{6}$$

$$R_s(t) = \exp(-(t/30239)^{0.86} - (t/26519)^{3.1032}) \tag{7}$$

2.6. Model Test

χ^2 test is a regular method to test a model when parameters are already known. In this case, according to the rang of sample, take 1×4000,2×4000,···,10×4000 to divide the number axis into 11 disjoint intervals, using accumulated running mileage of locomotives as observation samples, and assume that the number of occurrences of observation samples in the different interval obey the multinomial distribution. Then we can construct Pearson Statistics as Eq. (8), and take the significance level α=0.05 to test the two models.

$$\hat{\chi}^2 = \sum_{i=1}^{11}\frac{(n_i - np_i)^2}{np_i} \tag{8}$$

Where $\hat{\chi}^2$is the statistics of χ^2 test, n_i is the sample number within the i^{th} interval, and p_i is the accumulated probability of the given model within the i^{th} interval. Set the hypothesis as

H_0: $F_w(t) = 1 - \exp(-(t/35199)^{0.8754} - (t/34289)^{5.9318})$

and H_1: $F_w(t) \neq 1 - \exp(-(t/35199)^{0.8754} - (t/34289)^{5.9318})$

Get χ^2 test value in winter is: $\hat{\chi}^2 = 6.1959 < \hat{\chi}^2_{0.05}(10) = 18.3$. Then receive H_0 and reject H_1.

Similarly, the $\chi 2$ test value in summer is: $\hat{\chi}^2 = 10.95 < \hat{\chi}^2_{0.05}(10) = 18.3$, and also receive H_0 and reject H_1. It is thus verified that both the winter and summer models are feasible.

2.7. Analysis of Reliability Models

With Eqs. (6) and (7) as valid reliability models of diesel engine system for winter and summer, and their reliability distribution is illustrated in Fig. 3. Some analysis is introduced as following based on reliability functions and hazard rate functions.

As revealed by the plot, both the cumulative distribution curves keep the decreasing trend, which, however, still differs from what we expected. They were expected to decline more slowly at the beginning than in the middle, while it is reasonable according to the actual state of locomotives. It can be mainly attributed to the professional maintainability of local employees. For example, according to our statistics, in a preventively replace action to a low-temperature water pump, an incorrect assembling of the pump body caused the fracture of a pump shaft after running 1,867 km. Most malfunctions like this

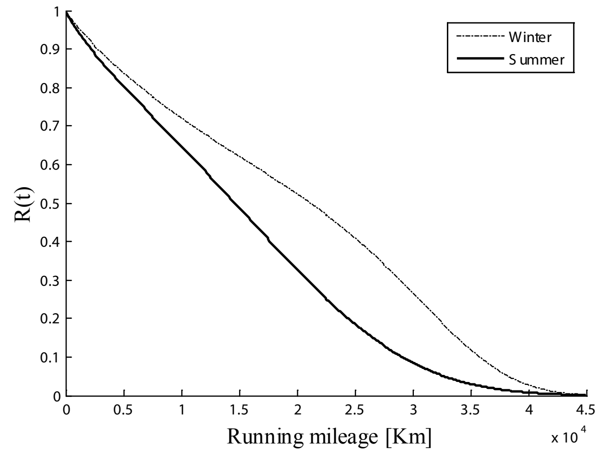


Fig. 3. Distribution of R (t) in Winter and Summer

example, which are caused by some incorrectly assembled or over maintained components, are main reasons that caused the reliability distribution curve quickly decreasing at the beginning, which does coincide with the authors' actual working experience and practice in repairing locomotives in this region. The current maintenance policy, briefly mentioned at the first section, is the main cause for the low reliability. Another typical example from our statistics is as follows: a a supply cam of the fifth cylinder was stripped off by severe wear-out due to insufficient maintenance, which ended in an engine failure after running over 3,216 km. Let us assume that the fault could be avoided if the maintenance action could be made in time, just to verify it can be found that it is inadvisable to extend the maintenance interval.

Let $r_w(t)$ and $r_s(t)$, separately denote the hazard rate of winter and summer, which relate to the reliability functions (6) and (7), and can get $r_w(t)$ and $r_s(t)$ as

$$r_w(t) = \frac{f_w(t)}{R_w(t)} = 2.49\times10^{-5}(t/35199)^{-0.1246} + 1.73\times10^{-4}(t/34289)^{4.9318} \tag{9}$$

$$r_s(t) = \frac{f_s(t)}{R_s(t)} = 2.84\times10^{-5}(t/30239)^{-0.14} + 1.17\times10^{-4}(t/26519)^{2.1032} \tag{10}$$

Two hazard rates are "bathtub-shaped" curves, by which we can crudely determine the maintenance interval. According to the monotonicity that the whole process can be clearly divided into three phases, which are marked as A, B(B' for winter) and C(C' for winter). As shown in fig. 4.

The phase A, about 0.2×10^4 km, is a running-in period after repairing. Our statistics reveals that the components which cause the

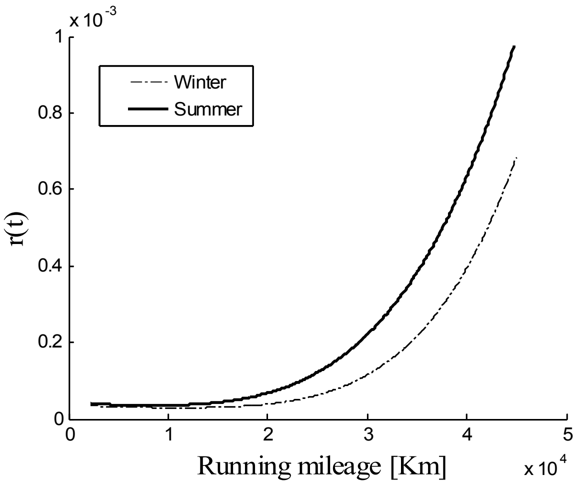


Fig. 4 Curves of Failure Rate Functions

engine failed more in summer than in winter are cooling pipelines of high/low-temperature system, and lubricating oil systems. The $B(B')$ phase stays about within 1.8×10^4 km in summer and 2.8×10^4 km in winter, which is termed as incidental period, whose failure rate goes close to a constant. The $C(C')$ phase is the worsening period, during which the failure rate goes up rapidly with the increase of the accumulated running mileage, this tendency being more distinct in summer than in winter.

The environment condition is known to be a main cause for higher failure rate in summer than in winter. From the failure rate value, we can see that diesel engine system fails great often in summer than in winter, mainly due to the summer temperature of over 40ºC that tends to make cooling systems operate under high load, unable to meet the cooling need of the entire locomotive in DF4B, and further cause the capability of lubricating oil becoming poor, and hence lead to wearing in worse in the incidental period. In contrast, it differs a lot in winter when low temperature (usually at −20 ºC) relieves the load of cooling systems by a great deal.

Therefore, judging by the hazard rate, it can be crudely regard that the maintenance interval in summer should be within 1.8×10^4 km, and in winter within 2.8×10^4 km. The current maintenance regulation say that the maintenance period does not less than 2.3×10^4 km, which may cause insufficient maintenance in summer and over maintenance in winter.

3. Maintenance Optimization

Based on the reliability models we got, and considered the requirements of railway enterprises, we comprehensively consider the effect of availability and economy to develop a maintenance policy. After that, the optimal intervals of PM period can be determined.

3.1. Optimization Based on Cost

To minimize the maintenance cost for the diesel engine system, the cost structure of the PM period is worth studying. Let C_f be the minimal repairing cost during the PM period, and $C_p(C_f>C_p)$ be the cost of preventively replacing components for maintenance. For a maintenance period, the expected cost E[C] and the expect cycle time E[T] can be calculated as follows[4]:

$$E[T] = \int_0^T tf(t)dt + T\int_T^\infty f(t)dt = \int_0^T R(t)dt \quad (11)$$

$$E[C] = C_f\int_0^T f(t)dt + C_p\int_T^\infty f(t)dt = C_fF(T) + C_pR(t) \quad (12)$$

The minimization of maintenance cost rate in a PM period can be represented as Eq. (13), where the numerator is equal to the expected total cost and the denominator equal to the expected total time.

$$\min: Z(T) = \frac{E[C]}{E[T]} = \frac{C_fF(T) + C_pR(T)}{\int_0^T R(t)dt} \quad (13)$$

Suppose that $\rho = \dfrac{C_f}{C_p}$

and then $\min: Z(T) = \dfrac{E[C]}{E[T]} = C_p\dfrac{\rho(1-R(T)) + R(T)}{\int_0^T R(t)dt}$

Substitute the reliability functions of both seasons for $R(T)$; as $\dfrac{dZ(T)}{dT} = 0$, we get

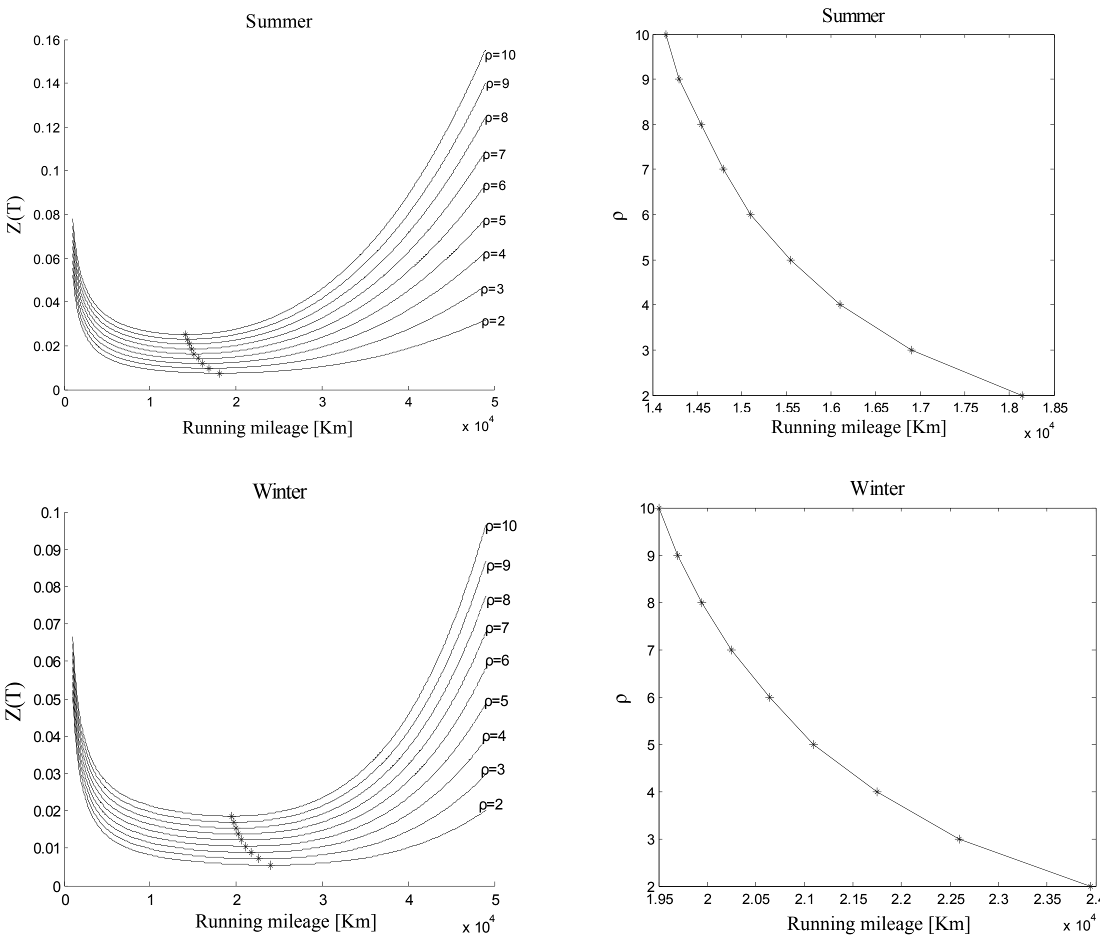


Fig. 5 ρ-(Z(T))

$$\left.\begin{aligned} & r(T)\int_0^T R(t)dt + R(T) = \frac{\rho}{\rho\text{-}1} \\ & \rho > 1 \end{aligned}\right\} \tag{14}$$

As indicated by (14), the limit of its right side is close to a constant when ρ is infinite, and the maintenance period T can be viewed as a constant when ρ is large enough. According to the history records, ρ is between 2 and 8. Herein, we set ρ=8, and then we get T_s=1.43×104km and T_w=1.97×10[4] km by Eq. (14), which are within the incidental period, and indicates apparently that there is a great gap between in summer and winter.

The optimization model based on maintenance cost is shown as (13), in which C_p and C_f are only relative to components. Based on the practice that ρ is between 2 and 8, we set ρ=2,3,…,8, and get the relationship curves for ρ and Z(T) as shown in Fig. 5.

Fig. 5 demonstrates tha ρ is approximately proportional to *minZ*(T). The striking variation of ideal maintenance period appears upon 1<ρ≤10. As in practice ρ is between 2 and 8, the expected maintenance period in summer is about 1.41~1.85×10[4] km, and that in winter is about 1.9~2.4×10[4] km. Fitting ρ and minZ(T) yields

$$E_w[C_\rho\ ] = C_p(-4.6\rho\ +183)E_w(T_\rho\ )\times 10^{-3} \tag{15}$$

$$E_s[C_\rho\ ] = C_p(-5.16\rho\ +241)E_s\,(T_\rho\ )\times 10^{-3} \tag{16}$$

Eqs. (15) and (16) can be considered as approximate experience formulas to determine the expected cost and the corresponding best maintenance period time.

3.2. Optimization Based on Availability

To maximize the efficiency of the diesel engine system, availability is also regarded as an index to optimize the maintenance period for maximum efficiency of the diesel engine system. Let T_f be the time for minimal repairing during the PM period, and T_p be the time for preventively replacing components in maintenance period. In this case, as the lifetime of diesel engine system is represented by the accumulated running mileage of locomotives, for description of the availability of diesel engines we need to convert T_p and T_f into corresponding equivalent kilometer t_p and t_f on the basis of the fact that locomotives finish transport assignment of 300 km every 8 hours. According to the maintenance regulations of railway enterprises, in general the maintenance assignment has to be finished within specified time. Therefore, letting T_p = 2 and T_f = 4 yields

thus $t_f = \frac{300}{8}\times T_f = \frac{300}{8}\times 4 = 150(Km)$

and $t_p = \frac{300}{8}\times T_p = \frac{300}{8}\times 2 = 75(Km)$

Let A be the availability and T the operation period. The maximum of E[A] is obtained as follows:

$$\max\!:\, E[A] = \frac{\int_0^T R(t)dt}{\int_0^T R(t)dt + T_pR(T) + T_fF(T)} \tag{17}$$

Set $0 < T_p < T_f < T$.

Substitute reliability functions $R(T)$ for summer and winter into equation (17), and according to $\frac{dE[A]}{dT} = 0$, we get

$$\begin{cases} r(T)\int_0^T R(t)dt + R(T) = \frac{T_f}{T_f - T_p} \\ 0 < T_p < T_f < T < 50000 \end{cases} \tag{18}$$

Solving Eq. (18) obtains T_s = 2.27×10[4] km and T_w = 2.83×10[4] km, both of which are longer than the optimization based on cost but run into worsening period. Therefore, it is unreasonable to optimize the maintenance period just by single factor.

3.3. Optimization Based on both Availability and Cost

For the purpose of obtaining the optimal period based on the fact of railway that centers availability of locomotives and considers the maintenance cost, and thus we take both the efficiency and the economy into consideration. Let $\frac{E(A)}{E(A)^*}$ represent the value function of availability of diesel engine system, $E(A)^*$ is the versus max(E(A)) in optimization based on availability. Let $\frac{Z(T)}{Z(T)^*}$ represent the value function of relative maintenance cost rate of diesel engines, $Z(T)^*$ is the *min*(Z(T)) in optimization based on maintenance cost. The final optimization model based on both availability and cost[15] can be shown as:

$$\min: ETC = -w_1\frac{E(A)}{E(A)^*} + w_2\frac{Z(T)}{Z(T)^*} \tag{19}$$

Where w_1 and w_2 are weighted values relative to decision-making tendency, and $w_1 \geq 0, w_2 \geq 0, w_1 + w_2 = 1$.

$$ETC = -w_1\frac{\int_0^T R(t)dt}{E(A)^*(\int_0^T R(t)dt + T_pR(T) + T_fF(T))} + w_2C_p\frac{\text{r}\ (1-R(T)) + R(T)}{Z(T)^*\int_0^T R(t)dt} \tag{20}$$

Considering that railway enterprises usually pay more attention to availability than to maintenance cost, so we adopt the optimization strategy that centers availability and also considers cost, set w_1=0.7, w_2=0.3, C_p=1000, and ρ=8. That $Z_w(T)^* = 0.017$, $Z_s(T)^* = 0.023$, $E(A)_w{}^* = 0.993$, and $E(A)_s{}^* = 0.990$ are known according to formula (10) and (13).

As $\frac{dETC}{dT} = 0$ we have

$$\begin{cases} -0.7\frac{75(r(T)\int_0^T R(t)dt + R(T))\text{-}150}{E(A)^*(\int_0^T R(t)dt + 75R(T) + 150F(T))^2} + 0.3\times 500\frac{7(r(T)\int_0^T R(t)dt + R(T)) - 8}{Z(T)^*(\int_0^T R(t)dt)^2} = 0 \\ 0 < T < 50000 \end{cases} \tag{21}$$

Thus, T_s = 1.575×10[4] km, and T_w = 2.125×10[4] km. The result indicates that the optimized maintenance period determined by addressing both availability and cost is within the incidental period and relative to w_1 and w_2. When w_1 = 0, the optimization is based on maintenance cost; when w_2 = 0, it is on availability. Both the values are relative to decision tendency.

3.4. Analysis of Optimization Results

According to the optimized results that the maintenance interval exist great gap between winter and summer, and the current maintenance period may cause insufficient maintenance in summer. The optimization result based on cost shows that the expected maintenance period in summer is 1.43×10[4] km, and that in winter is 1.97×10[4] km, both of them are within incidental period. Although the performance of diesel engine systems is stable in this period, the availability of locomotives is in a low. While the optimization result based on availability indicates that the maintenance period in summer is 2.27×10[4] km, and that in winter is 2.83×10[4] km, both of which are longer than the optimization result based on cost. While the result turns into worsening period, and the risk is increased. Therefore, it is thus clear that it is irrational to determine the maintenance period by only one factor. The third optimization result that centers the availability and considers the cost goes in accordance with the actual case. The decision tendency is determined by the weighted values. Taking into account the practice of railway enterprises, we set w_1=0.7, w_2=0.3, and get the optimal period in summer as 1.575×10⁴ km and that in winter as 2.125×10⁴ km, both of them are within incidental period.

Moreover, from the other side that the feasibility can be indicated from the comparison of maintenance cost and availability calculated by the formula (22) to (24) using the statistics, see table 1 and 2.

$$C_{real} = N_m C_f + N_p C_p \quad (22)$$

$$C_{op} = (C_f \int_0^T f(t)dt + C_p) N_p \quad (23)$$

$$E[A]_{real} = \frac{\sum_{i=1}^{N} t_i}{\sum_{i=1}^{N} t_i + N_m T_f + N_p T_p} \quad (24)$$

Where $E[A]$real and C_{real} and C_{op} separately denote the real availability and maintenance cost of the our statistics and the optimized maintenance cost, the optimized availability can be known by Eq. (18). N_m and N_p separately denote the whole times of minimum repair and preventive maintenance. The calculated results can be seen as Table 5.

As revealed by the above analysis, the maintenance period differs a lot between summer and winter, and the optimized results can improve the current maintenance period (≧2.3×10⁴ km) which maybe obviously decrease the locomotive's availability and increase maintenance cost, especially in summer. Meanwhile, it also indicates that the policy of extending the maintenance intervals adopted by certain railway enterprises is definitely undesirable.

Diesel locomotive is a complex electromechanical equipment, although malfunctions in diesel engine systems take up nearly 60% of those in the whole locomotive and the maintenance cost is also the majority, comprehensively analyzing maintenance period of diesel locomotives should consider diesel engine systems, running gears, electric apparatuses and brakes, and yet the maintenance period should not less than the results we got. Therefore, the result in this paper can be regarded as a reference to some railway enterprises.

Table 5. Comparison between real and optimized maintenance cost/availability

Items	Cost		Availability	
	Winter	Summer	Winter	Summer
Real result	¥625000	¥915000	0.9934	0.9921
Optimized result	¥533490	¥883910	0.9998	0.9997
Improved result	¥91510	¥31090	0.0065	0.0076

4. Conclusions

In the present study, the proposed double-fold competing risk models for winter and summer have been shown to give a good fit to the real data on diesel engines of locomotives of Urumqi Railway Bureau, and the optimized results indicate:

a) The maintenance period have great gap between winter and summer, railway enterprise should adopt different maintenance period in different season to avoid over maintenance.

b) The optimized maintenance can increase the availability and decrease cost more than the existing policy and can be regarded as a reference for Urumqi Railway Bureau and aroused their interest.

c) Obtained reliability models of diesel engine system can be used in grouped maintenance and performance improvement in diesel locomotives.

However, the diesel engine system is a complex electromechanical equipment, whose operating capability are influenced by many factors like outside environment, maintenance level of employee, operating level of engineer, track condition and so on. Therefore, further research on the topic is needed to address these factors using more accurate models than the ones proposed herein. Furthermore, this paper has also proposed a new method termed as ACPM for estimating multiple Weibull Competing Risk model parameters, which can get an objective fold and corresponding parameters rather than determined the fold in advance, at the same time, the method is easy to computer programming.

Acknowledgement

The research work is supported by the Foundation of the Key Laboratory of the Measurement and Control of Complex Engineering Systems, Ministry of Education of China (2010B001), and The Natural Science Foundation of Jiangsu Provence (50805023), and Hexa-Type Elites Peak Program of Jiangsu Province (2008144), and the Teaching and Research Fund for Excellent Young Teachers of Southeast University (200902), China.

Reference

1. Albarbar, A., F. Gub, and A.D. Ball, Diesel engine fuel injection monitoring using acoustic measurements and independent component analysis. Measurement 2010; 43(10): 1376–1386.
2. Bocchetti, D., et al., A competing risk model for the reliability of cylinder liners in marine Dieselengines. Reliability Engineering and System Safety 2009; 94(1): 1299–1307.
3. Fry, K.N. and BSc. Diesel locomotive a reliability improvement by system monitoring. Proc IMechE: Part F. 1995: 1-10.
4. J. Huang, C.R. Miller, and O.G. Okogbaa. Optimal Preventive-Replacement Intervals for the Weibull Life Distribution: Solutions & Applications. 1995 IEEE Proceedings Annual Reliability and Maintainability Symposium. Washington, DC, 1995: 370-373.
5. Jiang, R. and D.N.P. Murthy, Study of n-Fold Weibull Competing Risk Model. Mathematical and Computer Modelling 2003; 38(11): 1259–1273.
6. Jiang, R. and D.N.P. Murthy, Reliability modeling involving two Weibull distributions. Reliability Engineering and System Safety 1995; 47(3): 187–198.
7. Ling, D., H.-Z. Huang, and Y. Liu, A Method for Parameter Estimation of Mixed Weibull Distribution. IEEE Trans on reliability 2009; 1-4244-2509-9/09: 129–133.
8. Lingaitis, L.P., et al., Prediction Methodology of Durability of Locomotives Diesel Engines. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (2): 154–159.
9. M. Sarhan, A., David C. Hamilton, and B. Smith, Statistical analysis of competing risks models. Reliability Engineering and System Safety

2010; 95(9): 953–962.
10. Martinez, V. M., et al., Results and benefits of an oil analysis programme for railway locomotive diesel engines. Insight 2003; 45(6): 402–406.
11. Martorell, S., A. Sanchez, and V. Serradell, Age-dependent reliability model considering effects of maintenance and working conditions. Reliability Engineering and System Safety 1999; 64(1): 19–31.
12. Mazhar, M. I. and H.K. S. Kara, Remaining life estimation of used components in consumer products: Life cycle data analysis by Weibull and artificial neural networks. Journal of Operations Management 2007; 25(6): 1184–1193.
13. Park, C., Parameter Estimation of Incomplete Data in Competing Risks Using the EM Algorithm. IEEE Transactions on Reliability 2005; 54(2): 282–290.
14. Wei, D. I., Repair Period of Diesel Locomotives Based on Accumulative Damage Degree. China Railway Science 2009; 30(6): 104–107.(in Chinese).
15. Xia, T., et al., Optimal CBPM policy considering maintenance effects and environmental condition. Int J Adv Manuf Technol 2011; 56(9): 1181–1193.
16. Xu, R. and D. Wunsch, Survey of Clustering Algorithms. IEEE Transactions on Neural Networks 2005; 16(3): 645-678.

Prof. Zhisheng ZHANG, Ph.D.
Wenke GAO, Ph.D. Candidate
Yifan ZHOU, Ph.D.
School of Mechanical Engineering, Southeast University, Nanjing
China, 211189
E-mail: oldbc@seu.edu.cn; gaowk_best@163.com; yfzhou1981@163.com

Zhiqiang ZHANG, Eng.
Locomotive Sect of Urumqi Railway Bureau, Korla
China, 841000
E-mail: xj_zqzhang@163.com

Article citation info:
LIU Y, HUANG HZ, WANG Z, LI YF, ZHANG XL. Joint optimization of redundancy and maintenance staff allocation for multi-state series-parallel systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 312–318.

Yu LIU
Hong-Zhong HUANG
Zhonglai WANG
Yan-Feng LI
Xiao-Ling ZHANG

JOINT OPTIMIZATION OF REDUNDANCY AND MAINTENANCE STAFF ALLOCATION FOR MULTI-STATE SERIES-PARALLEL SYSTEMS

OPTYMALIZACJA ŁĄCZONA ALOKACJI NADMIAROWOŚCI ORAZ ALOKACJI PRACOWNIKÓW SŁUŻB UTRZYMANIA RUCHU W WIELOSTANOWYCH SYSTEMACH SZEREGOWO-RÓWNOLEGŁYCH

Multi-state system (MSS), as a kind of complex system consisting of elements with different performance levels, widely exists in engineering practices. In this paper, redundancy and maintenance staff allocation problems for repairable MSS with series-parallel configuration are considered simultaneously. The traditional redundancy allocation problem (RAP) for MSS always assumes that maintenance resources are unlimited. However in many practical situations, maintenance resources are limited due to the budget and/or time. To maximize the system availability under a certain demand, there are two feasible ways: (1) designing an optimal system configuration with available elements, and (2) allocating more maintenance staffs to reduce waiting time for repair. With the assistance of Markov queue model, the availabilities of identical version elements with the pre-assigned number of maintenance staffs can be evaluated. The universal generation function (UGF) is employed to assess the availability of entire MSS under a certain demand. Two optimization formulas considering the limited maintenance resources are proposed. One regards the limitation of maintenance resources as a constraint, and the other considers minimizing the total system cost including both the system elements and maintenance staff fees. The system redundancy and staffs allocation strategies are jointly optimized under required availability. A numerical case is presented to illustrate the efficiency of the proposed models. The Firefly Algorithm (FA), which is a recently developed metaheuristic optimization algorithm, is employed to seek the global optimal strategy.

Keywords: *multi-state series-parallel systems, redundancy allocation problem (RAP), maintenance staff allocation, queue theory, universal generation function (UGF), firefly algorithm (FA).*

Systemy wielostanowe (multi-state systems, MSS), stanowiące typ złożonych systemów zbudowanych z elementów o różnym poziomie wydajności, znajdują szerokie zastosowanie w praktyce inżynierskiej. W prezentowanej pracy podjęto rozważania łączące zagadnienia alokacji nadmiarowości oraz alokacji pracowników służb utrzymania ruchu w naprawialnych systemach MSS o konfiguracji szeregowo-równoległej. Tradycyjnie ujmowane zagadnienie alokacji nadmiarowości (redundancy allocation problem, RAP) w systemach MSS zawsze zakłada, że środki obsługi są nieograniczone. Jednakże w wielu sytuacjach praktycznych, środki obsługi mogą być ograniczone budżetem i/lub czasem. Istnieją dwa możliwe sposoby maksymalizacji gotowości systemu przy określonym zapotrzebowaniu użytkowników: (1) zaprojektowanie optymalnej konfiguracji systemu z wykorzystaniem dostępnych elementów oraz (2) alokowanie większej liczby pracowników obsługi w celu zmniejszenia czasu oczekiwania na naprawę. Dostępność jednakowych wersji elementów przy wcześniej określonej liczbie pracowników obsługi oceniano za pomocą modelu kolejek Markowa. Uniwersalną funkcję generacyjną (UGF) wykorzystano do oceny gotowości całego systemu MSS przy określonym zapotrzebowaniu. Zaproponowano dwa równania optymalizacyjne uwzględniające ograniczone środki obsługi. W jednym z nich ograniczoność środków obsługi potraktowano jako ograniczenie (constraint), natomiast drugie równanie dotyczyło minimalizacji całkowitych kosztów systemu włącznie z kosztami elementów systemu oraz płacą pracowników służb utrzymania ruchu. Strategie alokacji nadmiarowości systemu oraz alokacji pracowników poddano jednoczesnej optymalizacji z uwzględnieniem wymaganej gotowości. Wydajność proponowanych modeli zilustrowano przykładem numerycznym. Poszukiwania optymalnej strategii globalnej prowadzono przy pomocy niedawno opracowanego metaheurystycznego algorytmu optymalizacyjnego znanego jako algorytm świetlika (Firefly Algorithm, FA).

Słowa kluczowe: *wielostanowe systemy szeregowo równoległe, zagadnienie alokacji nadmiarowości (RAP), alokacja pracowników służb utrzymania ruchu, teoria kolejek, uniwersalna funkcja generacyjna (UFG), algorytm świetlika (FA).*

1. Introduction

Redundancy allocation problem (RAP), aims at providing redundancy at various stages of a system and determining an optimal system level configuration while considering the tradeoff between system reliability/performance and resources, has received increasing attention in reliability engineering as of late.

The reported works on RAP mainly focus on the problems of determining the optimal redundancy level for various stages, and/or selecting a set of components available in the market to maximize system reliability under some constraints, such as volume, weight and cost budget. They involve not only single/multi- objective optimization problem, but also binary/multi- state system configuration. For binary state systems, Coit et al. [4] introduced redundancy allocation optimization problem for binary state series-parallel systems. Coit and Konak [3] proposed a multiple weighted objectives heuristic algorithm to determine the optimal redundancy allocation solution for binary series-parallel systems. Other algorithms such as dynamic programming, integer programming, tabu search, and annealing simulation method, ant colony optimization approach were also proposed to solve the RAP in the context of binary state systems [6]. As an extension from binary state systems to MSSs, much research pays intensive attentions on RAP for MSSs. Lisnianski and Levitin [10] introduced RAP formulation for multi-state series-parallel systems, and the configuration of MSSs was determined by selecting an optimal set of components (elements) available in market. Tian et al. [20] proposed to jointly determine the optimal components state distribution of multi-state series parallel systems and its optimal redundancy level for each stage (defined as reliability-RAP). They formulated a multi-objective optimization problem, and physical programming was employed to solve the problem. Taboada et al. [19] proposed multi-objective multi-state genetic algorithm (MOMS-GA) to determine the optimal redundancy solution set under multi-criterions (cost, weight, availability etc.). Nourelfath et al. [17] developed an integrated model to jointly optimize the redundancy levels and imperfect preventive maintenance strategy for MSSs. A comprehensive survey of current advance in RAP can be referred to Ref.[6].

Most existing RAPs in literature assume that the maintenance resources are unlimited [18]. However, as stated in Ref.[1,12], the maintenance strategy suffers resources limitations in industrial applications, such as staffs, maintenance cost, and time, etc. Nourelfath and Dutuit [15] first proposed to solve the RAP considering repair policy. They discussed the issue when the number of repair teams is less than the number of reparable elements. A heuristic algorithm was applied to determine system configuration under availability constraint. Once a preliminary solution was found, stochastic Petri nets were used to model the different repair policies to evaluate the true system availability under limited repair teams. The optimization process was divided into two steps where the initial solution of the second step is conditional based on results from the first step. Nourelfath and Kadi [16] studied the same problem, dependencies resulting from maintenance teams sharing were taken into account. Universal generating function combined with Markov model was employed to calculate the element availability under limited maintenance staffs. They also employed a heuristic approach at the first step to optimize the system structure without considering the limitation of maintenance resources. To satisfy the constraint under limited maintenance staffs, system structure and maintenance resource allocation strategies are further optimized based on the solution obtained in the preceding step. Apparently, two-step optimizing process employed in the previous literature [15,16] cannot guarantee achieving the global optimal solution. Furthermore, maintenance activities were approximated via Markov queue model for each subsystem in Ref.[16]. Divergence may exist when the failure and repair rates are distinct between components within the same subsystem.

In this paper, the RAP of multi-state systems incorporating with staff allocation is studied. MSS is defined as a system being able to perform its task at different performance levels caused by degradation of components and parts in the system and/or the failure of some elements (more definition and applications of MSSs are referred to Ref.[10]). Different from the previous literature, elements available in market are chosen while the corresponding repair staffs are also allocated to maintain their availability. Markov queue model and universal generating function method are also employed to evaluate the availability [16] , and two optimization problems (PII and PIII) are proposed. The system configuration and the staff allocation strategy are optimized simultaneously through the firefly algorithm.

The remainder of this paper is organized as follows. The joint optimization problems are formulated in Section 2. In Section 3, the Markov queue model as well as the universal generating function method is presented to evaluate the availability of element and system with limited repair staffs. Section 3 briefly introduces the firefly algorithm and its implementation in the proposed optimization problems. A numerical case is given in Section 4 to verify the efficiency and effectiveness of the proposed models, and it is followed by a brief conclusion in Section 5.

2. Problem formulation

2.1. Definition of MSS

The MSS was primarily introduced in the middle of the 1970's by Murchland [14], and later discussed in Ref.[10]. According to the definition in Ref.[10], a system that possesses a finite number of performance rates is called an MSS. For example, if a flash memory chip fails in a computer system, the system can continues operate, but with derated memory capacity. This kind of system has a range of performance levels from its perfect functioning state to complete failure. There are many different situations in which a system should be treated as an MSS:

1) A system consisting of different units that have a cumulative performance effect on the entire system.
2) A system consisting of elements with variable performance due to deterioration (fatigue, partial failures etc.) and repairs actions.
3) A system with continuous performance deterioration is oftentimes simplified into the one with multiple discrete performance rates via state combination to reduce the computation burden [11-13].

In order to analyze the behavior of an MSS, one has to know the characteristics of its elements. Any system element j can have k_j different states corresponding to the performance rates, which is represented by the set:

$$\mathbf{g}_j = \{g_{j1}, g_{j2}, ..., g_{jk_j}\} , \tag{1}$$

where g_{ji} is the performance rate of element j in state i , $i \in \{1,2,...,k_j\}$.

The performance rate $G_j(t)$ of element j at any instant $t \geq 0$ is a random variable that takes its values from $\mathbf{g}_j : G_j(t) \in \mathbf{g}_j$. Therefore for the time interval $[0,T]$ where T is the MSS operation period, the performance rate of element j is defined as a stochastic process. The probabilities associated with different states of the system element j at any instant t can be represented by the set:

$$\mathbf{p}_j = \{p_{j1}(t), p_{j2}(t), ..., p_{jk_j}(t)\} , \tag{2}$$

where $p_{ji}(t)$ represents the probability that $G_j(t) = g_{ji}$. The state probabilities should satisfy the condition $\sum_{i=1}^{k_j} p_{ji}(t) = 1$ for any $t \geq 0$. Because the elements' states at any instant time t compose the complete group of mutually exclusive events.

The output performance of the entire MSS is defined as a stochastic process based on the system structure function:

$$G_s(t) = \phi(G_1(t), ..., G_N(t)) \tag{3}$$

where $G_i(t)$ is the performance stochastic process of the i^{th} element, and $\phi(\cdot)$ is system structure function. Thus, the probabilities associated with the different system state can be denoted by the set:

$$\mathbf{p(t)} = \{p_{s1}(t), p_{s2}(t), ..., p_{sK}(t)\} \tag{4}$$

where $p_{si}(t) = \Pr\{G_s(t) = g_{si}\}$, and K the number of possible system states, and g_{si} is corresponding performance at the i^{th} system state.

The system availability at time instant t for arbitrary demand W is given by:

$$A(t,W) = \Pr(G_s(t) \geq W) = \sum_{i=1}^{K} p_{si}(t) 1(F(g_{si}, W) \geq 0), \tag{5}$$

where $1(x)$ is unity function: $1(TRUE)=1$, $1(FALSE)=0$, and $F(g_{si}, W) = g_{si} - W$. If the demand is a random variable with M possible values, the availability of the MSS can be computed by:

$$A(t,W) = \sum_{i=1}^{M} q_i(t) \Pr(G_s(t) \geq W_i) = \sum_{i=1}^{M} q_i(t) \sum_{j=1}^{K} p_{sj}(t) 1(F(g_{sj}(t), W_i) \geq 0) \tag{6}$$

where W_i is possible user demand and $q_i(t)$ is corresponding probability at time t.

For a stationary system or a long time horizon, the instantaneous state probability at time t can be replaced by stationary state probability. Equations (5) and (6) will represent the stationary availability of a MSS.

2.2. Optimal design formulation

Before proposing optimization formulations, some basic assumptions are presented as follows:

(1) The MSS contains N_s subsystems connected in series. N_{is} versions of elements are available in market to be chosen for the i^{th} subsystem, and elements in the same subsystem are connected in parallel. Systems with this sort of configuration are called series-parallel MSSs.

(2) The elements in each subsystem are binary capacity elements. A binary capacity element i has two performance rates: nominal $g_{i1} \neq 0$ and $g_{i2} = 0$ for failure state.

(3) The elements available in market can be allocated to specified subsystems in a MSS. Thus, the MSS configuration is determined by choosing a set of element to assign to specified subsystems properly.

(4) The number of repair staffs is less than the number of elements in a MSS. One staff is just able to repair one version of element at a time. More elements of the same version exist in a MSS, more staffs are needed to keep a high element availability.

(5) The objective is to minimize the system cost under availability constraint.

The earlier reported works on RAP often ignore the limitation of repair staff (assumption 4). Thus, the optimization problem **PI** is formulated as (without considering constraints on weight, volume, etc.)

PI:

$$\begin{aligned} \min \quad & C_s = \sum_{i=1}^{N_s} \sum_{j=1}^{N_{is}} c_{ij} m_{ij} \\ s.t. \quad & A \geq A_0 \\ & m_{ij}^L \leq m_{ij} \leq m_{ij}^U \end{aligned} \tag{7}$$

where c_{ij} and m_{ij} represents the cost of the j^{th} version element and the corresponding number being used in the i^{th} subsystem, respectively. m_{ij}^L and m_{ij}^U are lower and upper bounds, and A_0 is the lower bound of availability constraint.

When considering the assumption 4, two types of optimization problems are proposed as follows. In the first type of problem, the staff cost is regarded as an additional constraint. The optimization formulation **PII** is given by

PII:

$$\begin{aligned} \min \quad & C_s = \sum_{i=1}^{N_s} \sum_{j=1}^{N_{is}} c_{ij} m_{ij} \\ s.t. \quad & A \geq A_0 \\ & C_{staff} = \sum_{i=1}^{N_s} \sum_{j=1}^{N_{is}} c'_{ij} n_{ij} \leq C_{s0} \\ & m_{ij}^L \leq m_{ij} \leq m_{ij}^U \\ & m_{ij} \geq n_{ij} \end{aligned} \tag{8}$$

where n_{ij} and c'_{ij} represent the number of repair staffs for the j^{th} version element and cost of per staff in the i^{th} subsystem, respectively.

In the second type of problem, the system construction cost incorporating with repair staff cost is treated as the objective. The problem **PIII** is formulated as

PIII:

$$\begin{aligned} \min \quad & C_s = \sum_{i=1}^{N_s} \sum_{j=1}^{N_{is}} \left(c_{ij} m_{ij} + c'_{ij} n_{ij}\right) \\ s.t. \quad & A \geq A_0 \\ & m_{ij}^L \leq m_{ij} \leq m_{ij}^U \\ & m_{ij} \geq n_{ij} \end{aligned} \tag{9}$$

2.3. Elements availability evaluation

As stated in previous section, we assume that every staff is just able to repair one version of element allocated in each subsystem. Suppose that, there exist m elements of version j in the i^{th} subsystem and n $(n \leq m)$ staffs available to repair failed elements as good as new. These elements are characterized by their identical failure rate λ_{ij} and repair rate μ_{ij}. With the assumption that an element just can be repaired by a staff at a time, it can be modeled through M/M/n queue theory as shown in Figure 1.

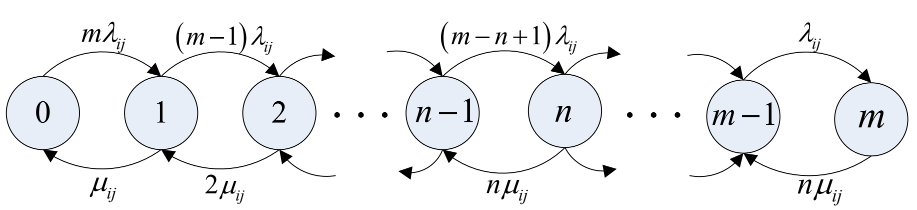


Fig. 1. The Markov diagram of M/M/n queue

The Markov transition intensity matrix is given by

$$Q=\begin{bmatrix} -m\lambda_{ij} & m\lambda_{ij} & & & & & & \\ \mu_{ij} & -[(m-1)\lambda_{ij}+\mu_{ij}] & (m-1)\lambda_{ij} & & & & & \\ & 2\mu_{ij} & -[(m-2)\lambda_{ij}+2\mu_{ij}] & (m-2)\lambda_{ij} & & & & \\ & & & \ddots & & & & \\ & & & n\mu_{ij} & -[(n-1)\lambda_{ij}+n\mu_{ij}] & (n-1)\lambda_{ij} & & \\ & & & & & \ddots & & \\ & & & & & n\mu_{ij} & -(\lambda_{ij}+n\mu_{ij}) & \lambda_{ij} \\ & & & & & & n\mu_{ij} & -n\mu_{ij} \end{bmatrix} \tag{10}$$

The stationary distribution can be derived by solving the corresponding Chapman-Kolmogorov equation:

$$\mathbf{P_{ij}}Q=\mathbf{0}\ , \tag{11}$$

where the vector $\mathbf{P_{ij}}=\{p_{ij}^0,p_{ij}^1,...,p_{ij}^m\}$ represented discrete probability distribution. The single p_{ij}^k is given by:

$$p_{ij}^k=\begin{cases}\dfrac{m!}{k!(m-k)!}(\dfrac{\lambda_{ij}}{\mu_{ij}})^k p_{ij}^0, & 0\le k\le n \\ \dfrac{1}{n!n^{k-n}}\dfrac{m!}{(m-k)!}(\dfrac{\lambda_{ij}}{\mu_{ij}})^k p_{ij}^0, & n<k<m \\ \dfrac{1}{n!}\dfrac{m}{n^{m-n}}(\dfrac{\lambda_{ij}}{\mu_{ij}})^m p_{ij}^0, & k=m\end{cases}\ , \tag{12}$$

where

$$p_{ij}^0=\left[\sum_{k=0}^{n-1}C_m^k(\frac{\lambda_{ij}}{\mu_{ij}})^k+\sum_{k=n}^{m}\frac{1}{n!}\frac{1}{n^{k-n}}\frac{m!}{(m-k)!}(\frac{\lambda_{ij}}{\mu_{ij}})^k\right]^{-1}, \tag{13}$$

and p_{ij}^k denotes the probability that k elements of version j in the i^{th} subsystem are available in a MSS. The rest $m-k$ elements are being repaired or waiting repair.

2.4. Universal generating function (UGF)

The UGF representing the probability mass function of a discrete random variable is defined by a polynomial form [7, 10, 21]. In the case of multi-state systems, UGF represents the random performance variable G_j of the elements and it is given by:

$$u_j(z)=\sum_{i=1}^{k_j}p_{ji}z^{g_{ji}}\ , \tag{14}$$

where the variable G_j has k_j possible values and $p_{ji}=\Pr\{G_j=g_{ji}\}$.

Therefore, in order to obtain the UGF of systems with arbitrary structure, one has to apply the composition operators $\otimes$ as follows recursively:

$$\begin{aligned} U_s(z)&=\otimes\{u_1(z),...,u_N(z)\} \\ &=\otimes\left\{\sum_{i_1=1}^{k_1}p_{1i_1}z^{g_{1i_1}},...,\sum_{i_N=1}^{k_N}p_{Ni_N}z^{g_{Ni_N}}\right\} \\ &=\sum_{i_1=1}^{k_1}...\sum_{i_N=1}^{k_N}\left(\prod_{j=1}^{N}p_{ji_j}z^{\phi(g_{1i_1},...,g_{Ni_N})}\right) \end{aligned}\ . \tag{15}$$

This polynomial represents all of the possible mutually exclusive combination of relating probabilities of each combination corresponding to the value of function $\phi(g_{1i_1},...,g_{Ni_N})$ which is determined by the system structure and performance rates combination property. For example, in the case of flow transmission system with two elements connected in series, one has:

$$\phi(G_1,G_2)=\min\{G_1,G_2\}\ , \tag{16}$$

and for the case where the two elements connected in parallel, the function is given by:

$$\phi(G_1,G_2)=G_1+G_2\ . \tag{17}$$

2.5. UGF of elements

In order to evaluate the reliability of a MSS with limited repair staffs, the UGF of elements availability is formulated at first. Assume that, in a flow transmission system, there are m_{ij} elements of version j and n_{ij} repairmen for these elements in the i^{th} subsystem. According to the Markov model in section 2.3, there exist totally $m_{ij}+1$ state for these elements. The UGF of the m_{ij} elements is formed as follows:

$$u_{ij}(z)=\sum_{k=0}^{m_{ij}}p_j^k z^{(m_{ij}-k)g_{j1}}\ , \tag{18}$$

where p_j^k is the probability that there are k elements of version j failed, which is achieved through the queue algorithm proposed in section 2.3. $(m_{ij}-k)g_{j1}$ is the corresponding performance rates at that state.

Thus, the UGF of each subsystem can be calculated through composition operation of the UGF of each version of elements. Then the UGF of MSS is achieved with iteratively operation as mentioned in section 2.4. Subsequently, the availability of the MSS under specified user demand can be determined according to Eqs. (5) and (6).

3. Firefly algorithm

Equations (7-9) are complicated non-linear programming problems. An exhaustive examination of all possible solutions is not realistic due to the limitation of computational time. Meta-heuristic algorithms such as Genetic Algorithm (GA), Tabu Search (TS), Simulated Annealing (SA) algorithm, and Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Firefly Algorithm (FA), Bat Algorithm(BA), are computationally efficient approaches to seek global optimal solution (or approximate optimal solution) in tough and complex optimization problems. The most attractive feature of these algorithms is that they are inspired by behaviors of biological systems and/or physical systems in nature. Also they possess intelligence to find global optimal solution even without derivative information. FA, a recently developed metaheuristic optimization algorithm, is employed in this paper to solve the proposed optimization problems since the superiority of FA over some other metaheuristic optimization algorithms was reported in Refs.[22-24]. The basic principle and its implementation in our problems are briefly introduced in the following sections.

3.1. Basic principle of firefly algorithm

Firefly algorithm inspired by the flashing behavior of fireflies was recently put forth by Yang [22-24]. The fundamental functions of flashing light of fireflies are to communicate (like attracting mating partners) and to attract potential prey. Inspired by this nature, the firefly algorithm was developed by idealizing some of the flashing characteristics of fireflies and representing each individual solution of optimization problem as a firefly in population. Three major idealized rules are [2, 5, 22]: (1) all fireflies in the population are unisex so that any individual firefly will be attracted at other fireflies; (2) for any pair of fireflies, the less bright one will move towards the brighter one. The attractiveness of a firefly is proportionally related to the brightness which decreases with increasing distance between two fireflies; (3) the brightness of a firefly is proportionally related to the value of objective function in the similar way to the fitness in genetic algorithm. The procedure of implementing the FA for a maximum optimization problem is summarized by the pseudo code shown in Figure 2 [5, 22-24].

Begin
Objective function $f(\boldsymbol{x})$, $\boldsymbol{x}=(x_1,\ldots,x_d)$
Define parameter of FA (light absorption coefficient γ *)*
Generate initial population of fireflies $\boldsymbol{x}_i$ (i=1.2, …n)
Determine brightness I_i *at* $\boldsymbol{x}_i$ *by objective function* $f(\boldsymbol{x}_i)$
Set counter t=1
while (t < *MaxIteration*)
for $i = 1 : n$ *all fireflies in population*
for $j = 1 : n$ *all fireflies in population*
If ($I_j > I_i$)
Move firefly i ($\boldsymbol{x}_i$) *towards j* ($\boldsymbol{x}_j$) *via Eq.(20)*
Evaluate value of objective function for firefly i ($\boldsymbol{x}_i$), *and update brightness* I_i *and attractiveness*.
end if
end for j
end for i
Rank fireflies by brightness and find the current best; $t = t + 1$
end while
Get the final best and postprocess results
End

Fig. 2. The pseudo-code of the FA

In the firefly algorithm, for simplicity, the attractiveness of a firefly is related to brightness I_i of the firefly which in turn is associated with the output of objective function $f(\mathbf{x}_i)$. For example, in the maximum optimization problems, the brightness I_i of the firefly i at location xi can be chosen as $I_i \infty f(\mathbf{x}_i)$. In nature, brightness decreases with the distance from its source, and light dims due to media. Therefore, the attractiveness β of one firefly to another is relative, and it should possess monotonically decreasing pattern with respect to the distance r_{ij} between firefly i and firefly j. In addition, the light absorption coefficient γ is also introduced to quantify the degree of light absorption. Several forms have been proposed to characterize the functional relationship of attractiveness β with respect to the distance r and light absorption coefficient γ. The following Gaussian form is used in the study:

$$\beta(r) = \beta_0 e^{-\gamma r^2} \tag{19}$$

where β_0 is the attractiveness at $r = 0$, and it equals to brightness. The light absorption coefficient γ can be either varied with respect to t or fixed [2, 22]. The distance r is defined as the Cartesian distance between a pair of fireflies $r_{ij} = \| \mathbf{x}_i - \mathbf{x}_j \|$. Firefly i will move towards firefly j if firefly j possesses a greater brightness than firefly i, and the movement is determined by [5]

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \beta_0 e^{-\gamma r^2} (\mathbf{x}_j^t - \mathbf{x}_i^t) + \alpha \boldsymbol{\varepsilon}_i \tag{20}$$

where $\mathbf{x}_i^t$ denotes the location of firefly i at the t^{th} iteration. The second term is due to the attraction from firefly j, and the last term is random movement. α is the randomization parameter and $\boldsymbol{\varepsilon}_i$ is a vector of random numbers drawn from a standard normal distribution representing the partial randomness of movement.

The iterative optimization process terminates once it meets some criterions, such as: 1) the number of iterations reaches the preset maximum value, and 2) the variance of average brightness in subsequent population is not more than ε_2, etc.

To apply the FA to a specific optimization problem, solution representation is an important procedure which must be defined. Penalty function approach can be employed to handle infeasible solutions.

3.2. Solution representation

For the optimization problem in Eqs. (8) and (9), the individual solution i is represented by the location $\mathbf{x}_i$ of firefly i as

$\mathbf{x}_i = \{\underbrace{x_1, x_2, ..., x_E}_{\text{redundancy number}}, \underbrace{x_{E+1}, ..., x_{2E}}_{\text{staff number}}\}$, where s_1 to s_E represent the number of redundancy for element of each version; and s_{E+1} to s_{2E} represent the number of staffs for element of each version. For example, a MSS consists of two subsystems. There exist two versions of elements in market for each subsystem. A specific individual solution $\mathbf{x}_i = \{\underbrace{1,2,0,3}_{part\,1}, \underbrace{1,1,0,2}_{part\,2}\}$ denotes that the MSS contains one version 1 element and two version 2 elements in subsystem 1, and three version 2 elements in subsystem 2. The repair staffs for each version of elements are 1, 1, 0, and 2, respectively. Since $\mathbf{x}_i$ only contain integers, the real numbers generated in the initial population and movements during iteration process have to be rounded off. In addition, on the ground that the number of staffs should be not greater than the number of redundancy of the corresponding element. Therefore in steps of initialization and movements, the number of staffs (part 2 of $\mathbf{x}_i$) of each individual solution $\mathbf{x}_i$ will be continually generated until its value is not greater than to the corresponding number of redundancy (part 1 of $\mathbf{x}_i$).

4. An illustrative case

Consider a flow transmission MSS consisting of four subsystems connected in series, and there are three versions of binary capacity element available in market for each subsystem. The parameters of these elements are tabulated in Table 1, as well as the cost for the corresponding repair staffs. The possible user demands at different levels are presented in Table 2 with the associated probabilities.

We solve the optimization problems **PI**, **PII**, and **PIII** under the same availability constraint $A_0 = 0.90$ and bounds ($m_{ij}^L = 0$, $m_{ij}^U = 5$). In our study, the firefly algorithm is performed to search a good solution in a computationally efficient manner. From our experimental tests, the values of parameters in FA are set as: $\alpha = 0.6$, $\gamma = 1.0$. The FA program is executed 10 times, and the optimal solution among

Table 1. The characters of element availability in market

	Ver. 1	Ver. 2	Ver. 3		Ver. 1	Ver. 2	Ver. 3
Subsystem 1				**Subsystem 3**			
Performance	120	85	65	Performance	130	100	75
λ_{1i}	0.0067	0.007	0.0065	λ_{3i}	0.0125	0.012	0.013
μ_{1i}	0.02	0.025	0.018	μ_{3i}	0.05	0.052	0.046
Cost ($)	1.5	1.2	0.9	Cost ($)	0.8	0.7	0.5
Staff Cost ($)	6.0	3.0	2.0	Staff Cost ($)	2.5	1.5	3.5
Subsystem 2				**Subsystem 4**			
Performance	100	95	65	Performance	125	95	65
λ_{2i}	0.0129	0.0135	0.012	λ_{4i}	0.00658	0.007	0.068
μ_{2i}	0.03	0.04	0.035	μ_{4i}	0.032	0.03	0.035
Cost ($)	3.5	2.5	2.0	Cost ($)	5.0	4.2	3.5
Staff Cost ($)	3.5	5.5	2.0	Staff Cost ($)	1.5	3.5	2.5

Table 2. The user demands

Demand (rate)	200	160	120	80
Probability	0.25	0.4	0.25	0.1

these results will be chosen as the final optimal result. The corresponding optimal solutions are tabulated in Tables 3, 4, and 5, respectively, where system configuration is listed in the column "Structure", and staff allocation strategy in the "Staff" column. For example, in Table 4, the solution "Subsys 1: 1-1-2-2-2" in the first row of the "Structure" column, denotes the subsystem 1 consists of 2 version 1 elements and 3 version 2 elements, while "1-2" in the "Staff" column of Table 4 represents to allocate 1 staff for version 1 elements and 2 for version 2.

In order to satisfy the availability constraint in problem **PI**, it requires at less 19 repair staffs with the cost equal to $51, and the total cost is equal to $88.7.

In problem **PII**, we considering that the number of repair staffs should be lower than 10, we set $c'_{ij} = 1$ and $C_{s0} = 10$, and the optimal results is presented in Table 4.

*Table 3. Optimal results for problem **PI***

Availability	Cost ($)	Structure
0.90011	37.7	Subsys 1: 1-1-2-2-2
		Subsys 2: 2-2-2-3-3-3
		Subsys 3: 1-2-2-2-3
		Subsys 4: 1-1-2

The staff cost in this case is equal to $34, and thus the total cost is equal to $78. If we assume the repair staffs are unlimited (at least 22 staffs) for this system configuration, the corresponding system availability is equal to 0.9468. It is indicated that the limited resource makes approximately 4.93% reduction in availability.

*Table 4. Optimal results for problem **PII***

Availability	Cost($)	Structure	Staff
0.9008	44	Subsys 1: 1-2-2-2-2	1-2
		Subsys 2: 2-2-2-2-3-3	2-1
		Subsys 3: 1-1-1-1-2-2-2	1-1
		Subsys 4: 1-1-2-2	1-1

The result for problem **PIII** is given in Table 5 and the total cost incorporating with staff cost are regarded as the objective to be minimized. It shows that the cost for the elements is equal to $44.8, which has nearly 1.82% increases compared to **PII** in the cost of elements. This result has approximately 11.15% reduction in total cost compared to **PII** under the same availability constraint. The result for **PII** ($78) is better than **PI** ($88.7). The optimal result in **PIII** has the least total cost ($69.3) when considering the staff cost jointly.

*Table 5. Optimal results for problem **PIII***

Availability	Total Cost ($)	Structure	Staff Cost ($)	Staff
0.90021	69.3	Subsys 1: 2-2-2-2-3-3	24.5	2-1
		Subsys 2: 2-2-3-3-3-3		1-2
		Subsys 3: 1-1-1-2-2-2-2		1-1
		Subsys 4: 1-1-1-1		2

5. Conclusions

In this paper, a joint optimization problem of redundancy and maintenance staff allocation for MSSs is proposed. The limited maintenance resource is a common issue in practices as emphasized by many researchers, but existing literature seldom discusses the RAP considering limited resources. Nourelfath et al. [15,16] first discussed the staff allocation in RAP through Petri nets and Markov model where maintenance strategies were considered within each module. As an extension of previous research, this paper allows the maintenance staff to be allocated for each version of elements. Without introducing any approximating approach in modeling as Ref. [16] , the M/M/n queue model and the UGF method are proposed to evaluate the availability of elements and system, respectively. Two optimization formulations concerning limitations from maintenance resources are introduced. The firefly algorithm, a recent developed metaheuristic algorithm, is employed to solve the resulting combinational optimization problems. Compared with the multi-step optimization approach proposed in Refs. [15, 16], the proposed approach which solve the optimization problem just in one step is more effective to achieve the global optimal solution. Also, as observed from our study, results from the proposed methods outperform the ones from traditional RAP without considering the limited maintenance staffs. However, this paper only studys the situation where the maintenance staffs allocate within the same subsystem and version of elements. Allocating maintenance staffs across the entire system is still an open research issue subjected to discussions in the future work.

Acknowledgments

The grant supports from the National Natural Science Foundation of China under contract number 71101017, the Research Fund for the Doctoral Program of Higher Education of China under contract number 20110185120014, and the Fundamental Research Funds for the Central Universities under contract number ZYGX2011J084 are greatly acknowledged.

References

1. Cassady C R, Pohl E A, Jr W P M. Selective maintenance modelling for industrial system. Journal of Quality in Maintenance Engineering, 2001; 7: 104-117.
2. Coelho L S, Bernert D L A, Mariani V C. A chaotic firefly algorithm applied to reliability-redundancy optimization. Proceedings of IEEE Congress on Evolutionary Computation, New Orleans, LA, USA, 2011, 517-521.
3. Coit D W, Konak A. Multiple weighted objectives heuristic for the redundancy allocation problem. IEEE Transactions on Reliability 2006; 55: 551-558.
4. Coit D W, Smith A E. Reliability optimization of series-parallel systems using a genetic algorithm. IEEE Transactions on Reliability 1996; 45: 254-260.
5. Gandomi A H, Yang X S, Alavi A H. Mixed variable structural optimization using firefly algorithm. Computers and Structures 2011; 89: 2325-2336.
6. Kuo W, Wan R. Recent advances in optimal reliability allocation. IEEE Transactions on Systems Man and Cybernetics Part A 2007; 37: 143-156.
7. Levitin G. The Universal Generating Function in Reliability Analysis and Optimization. London, Springer, 2005.
8. Levitin G, Lisnianski A. Joint redundancy and maintenance optimization for multi-state series-parallel systems. Reliability Engineering and System Safety 1999; 64: 33-42.
9. Lisnianski A, Estimation of boundary points for continuum-state system reliability measures. Reliability Engineering and System Safety 2001; 74: 81-88.
10. Lisnianski A, Levitin G. Multi-State System Reliability Assessment, Optimization, Application, Singapore, World Scientific, 2003.
11. Liu Y, Huang H Z, Reliability assessment for fuzzy multi-state system. International Journal of Systems Science 2010; 41: 365-379.
12. Liu Y, Huang H Z. Optimal selective maintenance strategy for multi-state systems under imperfect maintenance. IEEE Transactions on Reliability 2010; 59: 356-367.
13. Liu Y, Huang H Z, Levitin G, Reliability and performance assessment for fuzzy multi-state element. Proceedings of the Institution of Mechanical Engineers, Part O, Journal of Risk and Reliability 2008; 222: 675-686.
14. Murchland J. Fundamental concepts and relations for reliability analysis of multistate systems reliability and fault tree analysis. In Reliability and Fault Tree, Philadelphia: SIAM 1975; 581-618.
15. Nourelfath M, Dutuit Y. A combined approach to solve the redundancy optimization problem for multi-state systems under repair policies. Reliability Engineering and System Safety, 2004; 86: 205-213.
16. Nourelfath M, Kadi D A. Optimization of series-parallel multi-state systems under maintenance policies. Reliability Engineering and System Safety, 2007; 92: 1620-1626.
17. Nourelfath M, Châtelet E, Nahas N. Joint redundancy and imperfect preventive maintenance optimization for series–parallel multi-state degraded systems, Reliability Engineering and System Safety, 2012; 103: 51-60.
18. Nourelfath M, Fitouhi M, Machani, M. An integrated model for production and preventive maintenance planning in multi-state systems, IEEE Transactions on Reliability, 2010; 59: 496-506.
19. Taboada H A, Espiritu J F, Coit D W. MOMS-GA: A multi-objective multi-state genetic algorithm for system reliability optimization design problems. IEEE Transactions on Reliability, 2008; 57: 182-191.
20. Tian Z, Zuo M J, Huang H Z. Reliability-redundancy allocation for multi-state series-parallel system. IEEE Transactions on Reliability 2008; 57: 303-310.
21. Ushakov I A. Universal generating function. Soviet Journal of Computer and Systems Sciences 1986; 24: 118-129.
22. Yang X S. Nature-Inspired Metaheuristic Algorithm. Beckington, Luniver Press, 2008.
23. Yang X S. Firefly algorithms for multimodal optimization. Lecture Notes in Computer Science 2009; 5792: 169-178.
24. Yang X S. Engineering Optimization: An Introduction with Metaheuristic Applications, New York, Wiley, 2010.

Yu LIU, Ph.D., Assoc. Prof.
Prof. Hong-Zhong HUANG, Ph.D.
Zhonglai WANG, Ph.D., Assoc. Prof.
Yan-Feng LI, Ph.D. candidate
Xiao-Ling ZHANG, Ph.D.
School of Mechanical, Electronic, and Industrial Engineering
University of Electronic Science and Technology of China
No. 2006, Xiyuan Avenue, West Hi-Tech Zone, Chengdu, Sichuan, 611731
P. R. China
E-mail: hzhuang@uestc.edu.cn

Article citation info:
KAMIŃSKI Z, CZABAN J. Diagnosing of the agricultural tractor braking system within approval tests. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 319–326.

Zbigniew KAMIŃSKI
Jarosław CZABAN

DIAGNOSING OF THE AGRICULTURAL TRACTOR BRAKING SYSTEM WITHIN APPROVAL TESTS

DIAGNOSTYKA UKŁADU HAMULCOWEGO CIĄGNIKA ROLNICZEGO W RAMACH BADAŃ HOMOLOGACYJNYCH*

The paper describes the requirements for type approval testing of agricultural tractors with regard to braking, while taking into account the draft Regulation drawn up by the European Commission's Working Group on Agricultural Tractors (WGAT). A programme and methodology for testing the performance of the tractor brake system and its air brake system to supply and control the braking system of a towed vehicle are presented. Examples of diagnostic tests of a prototype tractor with hydraulically actuated brakes are given. These diagnostic tests may be wholly or partially used to develop a programme of qualification testing of tractors on production lines and a programme of periodic technical inspections of operated tractors.

***Keywords**: agricultural tractor, approval tests, diagnostics, brakes, air brake system.*

W pracy opisano wymagania dotyczące badań homologacyjnych ciągników rolniczych w zakresie hamowania, uwzględniające propozycje przepisów opracowywanych przez grupę roboczą Komisji Europejskiej ds. ciągników rolniczych (Working Group on Agricultural Tractors - WGAT). Przedstawiono program i metodykę badań skuteczności układu hamulcowego ciągnika oraz jego instalacji powietrznej do zasilania i sterowania hamulcami pojazdu ciągniętego. Zamieszczono przykłady testów diagnostycznych prototypowego ciągnika z hamulcami uruchamianymi hydraulicznie. Opisane testy diagnostyczne mogą być w całości lub częściowo wykorzystane do opracowania programu badań kwalifikacyjnych ciągników na liniach produkcyjnych oraz programu okresowych badań technicznych ciągników eksploatowanych.

***Słowa kluczowe**: ciągnik rolniczy, badania homologacyjne, diagnostyka, hamulce, instalacja pneumatyczna hamulcowa.*

1. Introduction

Type approval tests are an integral part of the official procedure that determines a vehicle being approved to be allowed to move on public roads. Type approval tests are only conducted by authorized testing establishments following the testing methods specified in relevant regulations. The aim of the diagnostic tests of both newly manufactured or operated vehicles is to verify and assess the technical conditions of subassemblies that have a significant impact on the technical performance of the vehicle in the light of the requirements imposed on them.

In view of the road safety, the braking systems of agricultural vehicles must meet a number of requirements for, among other things, braking efficiency, the follow-up action during slow braking, and a high speed of action during sudden braking (a response time less than or equal to 0.6 s). Tests carried out in the UK on tractor units with trailers equipped with hydraulic braking systems [14] showed that as many as 90% of the operated trailers did not attain the required value of the braking rate (z=0.25); after making the necessary adjustment, the fraction of trailers failing to meet the braking rate fell to 40%. Particularly hazardous is the operation of assemblies of incompatible vehicles due to the different efficiency of action of their respective braking systems (braking asynchrony) [12]. The operation of a high-speed modern agricultural tractor with high-efficiency brakes coupled with a low braking-efficiency trailer will, on the one hand, lead to the accelerated wear and premature damage of the trailer's braking system and, on the other hand, cause overloading, rapid wear and possible damage of the tractor's braking system [14]. The incompatibility of the braking systems of vehicles in tractor units, resulting in jack-knifing or skidding during braking, was the cause of about 9.7% fatal road accidents involving agricultural vehicles in the UK in the years 1999-2004 [4].

In Poland, air braking systems are commonly used in agricultural trailers. Therefore, agricultural tractors used for transport, regardless of the design of their own brakes (hydraulic – Fig. 1, mechanical, or pneumatic), are furnished with pneumatic systems intended for braking the towed trailers. Older tractors are equipped with single-line air braking systems, while in newer ones, there are usually combined systems that allow the use of both older trailers furnished with single-line air braking systems and new trailers that have dual-line air braking systems. The recommendation to use dual-line systems and increase the braking rate of tractors to 0.45 was introduced in 2004 within the amendment of the regulations on the approval of agricultural tractors for braking [6], [13], because of increasing the maximum permissible driving speed to 40 km/h for tractors of categories T1–T4 and above 40 km/h for tractors of category T5. In practice it happens that T5 category tractors move at a speed of around 60–70 km/h. The unsatisfactory state of implementation of the new regulations, the consent on continuing the manufacture of trailers with obsolete single-line braking systems and the operation of agricultural vehicles with inoperative or incompatible braking systems or even with no

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

braking system, as is the case for forestry trailers [11], all constitute a threat to the safety of road traffic involving those vehicles. Each year in Poland, of about 900 accidents involving agricultural vehicles, 90-100 have tragic consequences [5]. Therefore some [2] even call for the enforcement of upgrading the air braking systems of agricultural vehicles and trailers and introducing the inspection of braking systems to be carried out by Vehicle Inspection Stations.

For several years now, work has been continued in the European Union on the comprehensive proposal for the unification of technical regulations regarding the braking of agricultural and forestry vehicles modelled on Regulation 13 of the European Economic Commission for motor vehicles [8]. The recent proposals by the Working Group on Agricultural Tractors (WGAT) [7] assume, among other things, increasing the braking rate $z \geq 0.5$ for agricultural tractors moving at a speed of above 30 km/h and trailers and towed agricultural machinery with a Total Technically Permissible Mass (TTPM) of over 1.5 tons (trailers of categories R2, R3, R4; and towed agricultural machines of category S2). It is also proposed, similarly as for an assembly of road vehicles, to introduce compatibility corridors, that is the permissible variations in the braking rate value as a function of pressure at the coupling head connecting the tractor's and the trailer's control line for tractors moving at a speed of above 30 km/h and trailers with an TTPM of above 3.5 tons (categories R3 and R4) and towed agricultural machines of category S2.

A key issue for the improvement of the quality and safety of transport means for the agricultural and forestry sectors, in the light of the above-mentioned technical and legal status of their braking systems, seems to be to improve the methods and tools for diagnosing the braking systems of agricultural vehicles within the framework of approval and qualification tests and periodical technical inspections. The present paper describes the methodology for diagnosing agricultural tractor braking systems within approval tests following the programme [9] taking into account the requirements of the Regulation for agricultural and forestry vehicles as developed by the WGAT [7] and the Regulation 13 of the European Economic Commission [8] for motor vehicles. Example results of tests on a tractor with hydraulically actuated brakes are provided.

2. Requirements of braking systems

The current requirements of tractor braking systems contained in applicable [6], [8], [10], [13] and proposed [7] regulations can be grouped as shown in the schematic in Fig. 1, similarly as for motor vehicles [1].

Groups I and II of requirements apply to the supply unit (compressor, pressure regulator, compressed air reservoirs), whereas group IV applies to the air brake system control unit. The requirements of group III are applicable to the tractor's service and parking braking systems.

2.1. The energy source

The criterion for the performance assessment of the compressor (**I** – in Fig. 1) in the supply unit is assumed as the time of filling the tractor's reservoir and the substitute reservoir connected to the supply coupling head to a pressure value p, as specified by the manufacturer, at which the vehicle attains the braking efficiency prescribed for the service brakes. The capacity of the substitute reservoir imitating the

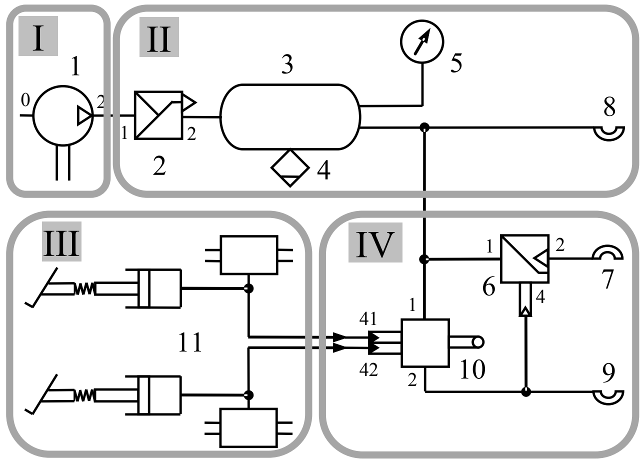


Fig. 1. Schematic diagram illustrating the division of a tractor braking system according to the requirements: I – energy source, II – distribution of energy and assurance of its required level, III – tractor braking system, IV – trailer brake control system (a combination of the tractor and trailer pneumatic systems), 1 – compressor, 2 – unloader valve (pressure regulator), 3 – air reservoir, 4 – drain valve, 5 – manometer, 6 – inversion trailer control brake valve, 7 – "single-line" coupling head (black), 8 – supply coupling head (red), 9 – control coupling head (yellow), 10 – trailer brake control valve, 11 – tractor hydraulic brake system

capacity of the trailer supply system is calculated from the relationship:

$$V = 20 \cdot R / p_{\max} \quad (1)$$

where: p_{max} – maximum value of pressure controlled in the system [bar], R – permissible maximum load on the all trailer axles [t].

The times t_1 and t_2 of filling the reservoirs to 65% and 100% of the p pressure value, respectively, as given in Table 1, should be measured during the compressor operating at the maximum power or the maximum rotational speed of the combustion engine. The value of pressure p is normally equal to the value of minimum controlled pressure $p_{\min}$. If additional reservoirs intended for accumulating energy needed for actuating devices not belonging to the brake system are installed in the tractor, and their capacity does not exceed the 20% of the total capacity of the braking system reservoirs, then, in accordance with the new regulations [7], the time t_3 needed for filling the reservoir most unfavourably situated relative to the compressor must also be determined.

2.2. Distribution of energy and assuring its required level

Maintaining the necessary energy level in the supply unit (**II** – in Fig. 1) of the tractor's pneumatic system depends on the capacity of the compressed air reservoir, pressure control parameters (i.e. the values of minimum and maximum controlled pressure, p_{min} and p_{max}) and, in high-pressure systems, also on the unloader valve.

With the compressor not operating and at the initial reservoir pressure equal to p, after eight full actuations of the service brake the value of reservoir pressure should not be lower than the value necessary for achieving the efficiency of the emergency brake. The location of control pressure measurement is the 0.5 dm³ capacity compressed air reservoir connected to the control line and vented prior to each

Table 1. Time of filling the supply unit during testing of the compressor [7], [8], [10]

Pneumatic supply system option	Filling time [min]		
	t_1 at 65% p	t_2 at p	t_3 at p
The system of a vehicle designed for towing a trailer	6	9	11
The system of a vehicle not designed for towing a trailer	3	6	8

successive braking; the supply line should be blanked off. In tractors designed for towing trailers, the pressure in the control line after 8 consecutive braking events should have a value not lower than half the value obtained after the first brake aplication.

For single-line systems, the reservoir capacity should be selected such that after the complete one-off braking and releasing cycle a pressure reduction from the minimum value be not greater than 0.5 bar [10]. The pressure is measured at the end of the 2.5 m-long 13 mm-internal diameter line connected to the tractor's supply and control coupling head (a line simulating the capacity of the trailer's supply & control line).

2.3. The tractor brake system

Brakes (**III** – in Fig.1) have a decisive effect on the safety of driving under normal operation conditions, therefore special requirements are laid on them. These relate to the braking performance of both cold and hot brakes, acting speed, the compatibility of the braking systems of vehicles making up a road unit, and the design of the system.

The efficiency prescribed for a braking system should be determined based on the stopping distance s_z or the mean value of fully developed deceleration d_m:

$$s_z = \frac{v}{3.6}\left(t_o + \frac{1}{2}t_n\right) + \frac{1}{2}\frac{(v/3.6)^2}{d_m} \quad [m] \tag{2}$$

where: v – initial speed [km/h]; t_o – brake actuation time referred to as the braking system delay time [s]; t_n – braking deceleration increase time [s], d_m – fully developed braking deceleration [m/s²].

The requirements for cold tractor brakes (test of type 0) are summarized in Table 2. The required decelerations for the emergency brakes are given in parentheses.

Table 2. Requirements for the efficiency of tractor service and emergency brakes in tests with cold brakes – (the 0 type test); the recommended method of checking is highlighted with the grey background

Vehicle category	s_z [m]	d_m [m/s²]	z= d_m/g	Conditions
Applicable requirements [10], [13]				
T1÷T5	≤0.15 v + v²/116	≥4.5	≥0.45	
Proposed by the WGAT [7]				
T1÷T4	≤0.15 v + v²/116	≥4.5 (1.5)	≥0.45	vmax ≤ 30 km/h
T1÷T4, T5	≤ 0.15 v + v²/130	≥ 5.0 (2.2)	≥0.50	vmax > 30 km/h

The initial velocity v may not be lower than 98% of the velocity prescribed for given tests (the maximum velocity). The mean fully developed deceleration dm should be calculated as the average deceleration related to the distance in the range from v_b to v_e according to the following formula:

$$d_m = \frac{v_b^2 - v_e^2}{25,92(s_e - s_b)}[m/s^2] \tag{3}$$

where: v_b – vehicle velocity corresponding to 0.8 v [km/h]; v_e – vehicle velocity corresponding to 0.1 v [km/h], s_b – distance covered between v and v_b [m]; s_e – distance covered between v and v_e [m].

Next, the efficiency of vehicle braking with hot brakes is checked. This is the so called fade test (type I). After the type I fade test with hot brakes, the efficiency of the service brakes should not be less than 75% of the prescribed efficiency and not less than 60% of the value measured in the 0 type test (for cold brakes). The basic differences between the regulations being currently in force and the proposed regulations result from the brake heating mode and conditions. In accordance with the applicable requirements [13], heating of brakes should be carried out in such a manner that the energy lost in the brakes corresponds to the energy needed for maintaining the vehicle velocity at a level of 80%±5% of the maximum velocity per 1 km of a 10% slope road stretch.

In the proposed regulations [7], the fade test of hot tractor brakes should be performed with repeated braking through consecutive applications and releases of the brakes. The number of braking and releasing cycles n and the testing conditions are given in Table 3. The force applied to the control should be selected so that the mean fully developed deceleration value equal to 3 m/s² could be achieved with the first braking. This force should be maintained in all subsequent brakings.

Table 3. Conditions of service brake efficiency testing (the I type test with hot brakes)

Vehicle category	Conditions			
	v_1 [km/h]	v_2 [km/h]	Δt [s]	n
T	80% v_{max}	½ v_1	60	20

The respective quantities in Table 3 are denoted as follows: v_1 – velocity at the start of braking; v_2 – velocity at the end of braking; n – number of brakings; Δ_t – duration of the braking cycle between the start of a braking and the start of the next braking.

The tractor parking brake system should assure the tractor to be immobilized on both a declivity and an acclivity with an 18% slope (for category T4.3 the acclivity slope value is 40%). In the case of a vehicle unit with an unbraked trailer, the tractor should hold the vehicle unit immovable on both a declivity and an acclivity with a 12% slope. As per the applicable regulations, the mass of an unbraked vehicle may not be greater that the mass of the tractor and may not in any case exceed 3 tons. In the proposed regulations, the maximum mass of a drawing vehicle-towed unbraked vehicle unit should not exceed the maximum permissible mass of the loaded tractor multiplied by the quotient of the prescribed maximum stopping distance by the stopping distance as determined in the 0 type test [7]:

$$P_C \le P_M \frac{s_p}{s_a} \tag{4}$$

where: P_C – maximum mass of the combination of the tractor and the un-braked towed vehicle, as declared by the tractor manufacturer, P_M – maximum mass of the laden tractor, s_p – prescribed stopping distance, s_a – achieved stopping distance measured during type 0 test (the tractor laden to its maximum mass P_M). In any case, the total load on all axes of the towed unbraked vehicle should not exceed 3.5 tons.

The emergency brake system of a tractor should reduce the speed of the vehicle until its stop with deceleration equal to at least 1.5 m/s² when v_{max}≤30 km/h, and 2.2 m/s² when v_{max}>30 km/h. The efficiency tests of emergency brakes should be done by simulating a failure of the service brake system.

Braking systems are required to show a follow-up action in slow breaking and a high-speed action in sudden braking. The notion of follow-up action is understood as the braking system ability to maintain the proportional relationship between the input signal change and the output signal change under steady state conditions. This means that the pressure in the actuator chambers should change proportionally to the displacement of the brake pedal.

During sudden braking, the action speed of the braking system of the tractor and the trailer should be such that the prescribed service brake efficiency be achieved in a time not longer than 0.6 s. In hydraulic braking systems, this condition is considered satisfied, if the braking deceleration or the brake fluid pressure in the most remote actuator attains a value corresponding to the prescribed braking efficiency. In sudden braking, a pedal force increase time of 0.2 s is assumed.

2.4. Trailer brake control

In dual-line systems, the connection between the tractor pneumatic system and the trailer braking system is done using two lines, of which one is intended for supplying and the other for controlling the trailer braking system. To assure the exchangeability of tractors and trailers, the values of pressure in the both lines have been unified. With the full application of the service brake, the pressure values should lie in the range from 6.5 to 8.5 bar. In single-line systems, the pressure at the coupling head should range from 5.8 to 6.3 bar.

The control of the trailer brakes (**IV** – in Fig. 1) should be effected exclusively with the simultaneous application of the drawing vehicle's brakes, therefore the required times of pressure variations in the couplig head of the drawing vehicle's control line are specified. In a single-line system, the time elapsing from the start of pressing the brake pedal until the moment when the pressure in the control line (2.5 m long and 13 mm in internal diameter) connected to the tractor's coupling head decreases to 90% of the minimum value should not exceed 0.2 s, and when the pressure decreases to 25% of this value, it should not exceed 0.4 s with a full pedal application time of 0.2 s [10]. In the proposed new regulations [7], there is no requirement for testing for the response time of the control unit for dual-line systems. To evaluate the action speed of this unit, the provisions of EEC Regulation 13 [8] can be used. In dual-line systems, the time of pressure build-up in the control line up to 10% of the asymptotic pressure should not exceed 0.2 s, and up to 75%, 0.4 s. The requirements for the response time of the tractor control unit are illustrated in Fig. 2.

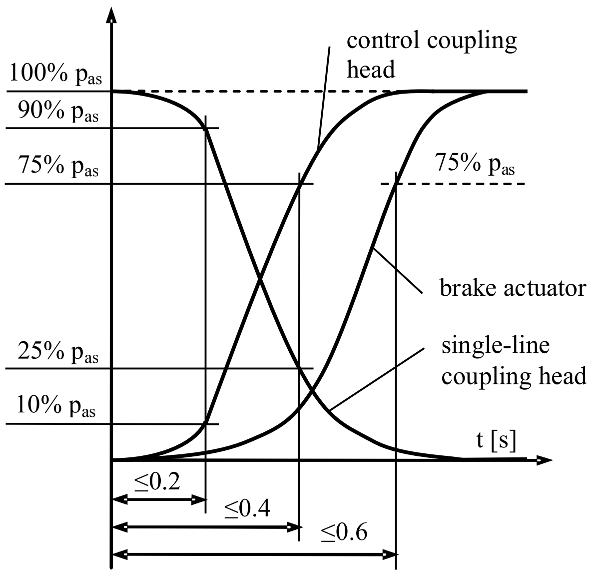


Fig.2. Variation in pressure as a function of time and the required response times of the tractor control unit

In addition to the requirements for action speed, there are also requirements for the compatibility of the braking systems of a tractor unit's vehicles [7]. Braking of the trailer only in an assembly of vehicles is not permissible. The proper synchrony of action of the both vehicles' brakes should prevent the loss of running stability that would lead to a folding of the unit. The condition for the proper braking synchronism is the selection of trailer braking intensity as a function of control line pressure such that a slight tensile force arises in the towing attachment in the first braking phase, which will maintain the alignment of the tractor and the trailer.

2.5. General technical requirements

Aside from the requirements that determine the roadworthiness of a vehicle, there are specific requirements for the tightness of the system and the reliability and durability of assemblies and parts under varying conditions of operation of pneumatic braking systems.

The leakage of the system is checked by measuring the pressure drop in the system after a specified time. According to [10], the reservoir pressure drop from the minimum controlled pressure value should not exceed 2% for 10 minutes. Much less strict requirements are used by the Wabco company, which assume as permissible a 5% pressure drop after 3 minutes in brake actuators pressurized initially up to half the maximum controlled pressure value [1].

3. Testing methodology

The primary goal of the undertaken diagnostic tests was to try out the methods and procedures of the developed testing programme [9] including the efficiency testing of the brakes of a tractor and testing of its air braking system. The air braking system testing programme included the verification of:

- air system leakage,
- unloader valve operating range,
- coupling head pressure values,
- compressed air reservoir capacity,
- air compressor capacity,
- control unit response time.

The tests were carried out for a Pronar 81.6 KM prototype tractor equipped with hydraulically actuated brakes. The tests of the tractor's pneumatic braking system included the recording of the time variations of the brake pedal force and air pressure at selected points in the system. The measuring method, testing conditions and the requirements of the subassembly tested are described in detail in reference [9]. The measuring system consisted of:

- a Pentium III PC,
- an input-output adapter,
- an MC1212 measuring card by Senga with a resolution of 12 bits and a processing accuracy of 0.02% FSR ±1 LSB,
- a CL 23 force strain gauge by ZEPWN complete with a converter, with a measuring range of 0÷1 kN and accuracy class 0.1,
- MBS 32 pressure transducers by Danfos, with a measuring range of 0÷10 bar, accuracy class 0.3, DMT-21 digital revolution counter (for measuring the engine rotational speed) with a measuring range of 0÷9999, accuracy class 0.2, and
- the „MC1212" the program for the recording and acquisition of measurements.

The efficiency tests of the tractor's main brakes included the recording of the distance, deceleration and the brake pedal force during the braking process. For recording, a measuring system (Fig. 3) described in detail in reference [3] was used, which was made up of:

- a portable computer,
- an LTC1286 measuring module,
- a CL 23 force strain gauge by ZEPWN complete with a converter, with a measuring range of 0÷1kN, accuracy class 0.1,
- a fifth wheel with a rotary-pulse converter,
- an electronic decelometer with a measuring range of ±2g,
- Holux GPSlim 236 GPS receiver,
- the „POMIAR1286" program for recording data during braking.

The use of the LTC1286 measuring module including a 12-bit LTC1286 A/C converter and an HCF4051 multiplexer enabled the frequency of measurements from the analog outputs to be increased to 10 kHz. In addition, the module had 8 digital inputs and 8 digital outputs relying on HCF4051 and MC4094circuits. Communication with

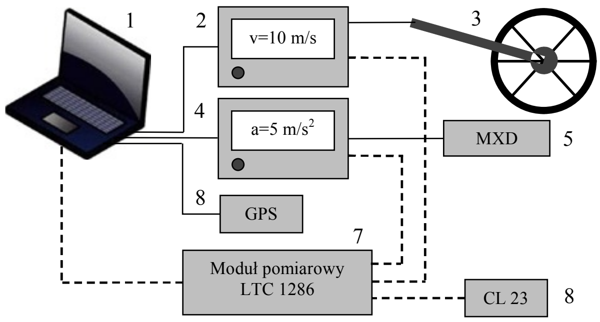


Fig. 3. Schematic of the measuring system for testing the vehicle braking process: 1 – portable computer, 2 – "fifth wheel" controller, 3 - fifth wheel, 4 - electronic decelometer, 5 - MXD7202 acceleration sensor, 6 – Holux GPSlim 236 GPS receiver, 7 – LTC1286 measuring module, 8 – CL23 brake pedal pressure sensor

the portable computer was effected through the LPT port or the USB connection using the SPI synchronous data transmission interface.

For recording the braking distance, the fifth wheel was used, in which the pulses generated by the disc were recorded by a TCST1103 optical sensor (Fig. 4-a) and then counted by a microcontroller relying on the ATMega 8 circuit and programmed in the Bascom language. Based on the counted pulses, the distance covered by the vehicle and the vehicle instantaneous velocity and acceleration were determined. The microcontroller system was equipped with an RS232 interface for its calibration (storing of the wheel rolling radius value and the number of disc pulses in the EEPROM memory) and the transmission of measurement data to the portable computer.

The braking deceleration was determined by an indirect method through the differentiation of the signal received from the fifth wheel and by a direct method using accelerometric sensors as shown in Fig. 4-b. For their construction, MXD7202 acceleration sensors manufactured by MEMSIC were used, which were connected via a signal amplifier to the microcontroller furnished with an RS232 interface for communication with the portable computer. Circuit diagrams for the measuring module and the fifth wheel and decelometer microcontrollers are provided in reference [3]. The „Pomiar1286" program written in the environment was used for handling the measuring system.

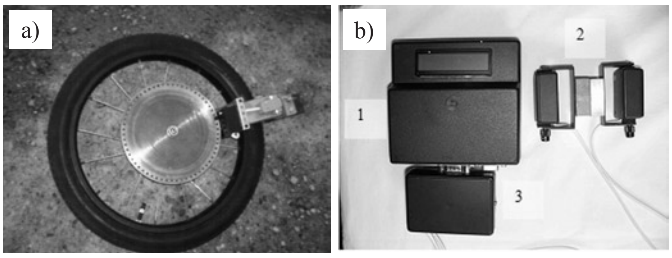


Fig. 4. Devices for measuring the braking distance and braking deceleration: a – fifth wheel pulse disc with an optical sensor, b – decelometer components: 1 – microcontroller, 2 – measuring sensors, 3 – signal amplifier

4. Examples of diagnostic tests

4.1. The air system leakage test

The leakage test involved recording of the 10 minute drop in pressure $p_z(t)$ at the supply coupling head of a dual-line system from the initial value equal to the minimum controlled pressure value p_{min}. Should a leak be detected (Fig. 5), it would be necessary to locate and remove the leakage before proceeding with subsequent tests.

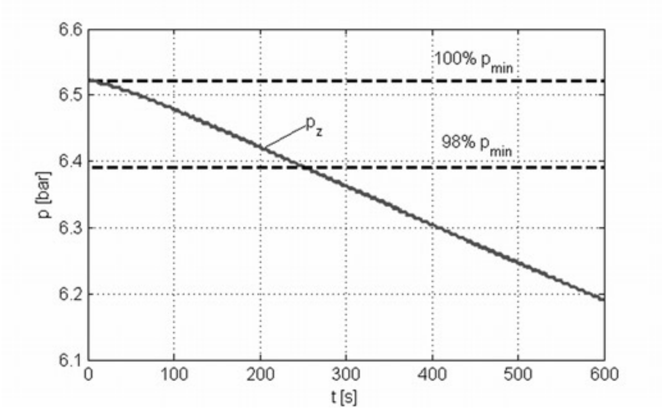


Fig. 5. Testing the air system leakage based on the variation of pressure p_z at the supply coupling head

4.2. The unloader valve operation test

The unloader valve operation test was performed with the operating engine by recording changes in pressure in the tractor's compressed air reservoir which was vented through the drain valve. A 385±5 cm³ capacity reservoir was connected to the control coupling head of the dual-line system, which was an equivalent for the capacity of the 2.5 m-long 13 mm-diameter trailer supply line. Based the recorded variation of pressure p_{zb} (Fig.6), the values of the unloader valve switch-on pressure p_{min}=6.40÷6.47 (6.5÷6.8 bar acc. to the specification) and switch-off pressure p_{max}=7.85÷7.88 bar (8±0.2 bar acc. to the specification) were determined, at a confidence level of 95%. The tests demonstrated the need for controlling the switch-on pressure. From the tests, incorrect Visteon 51 10 018 inloader valve switch-on pressure values, not conforming to the specification, were found. However, no irregularities in the cyclic operation of the unloader valve were identified.

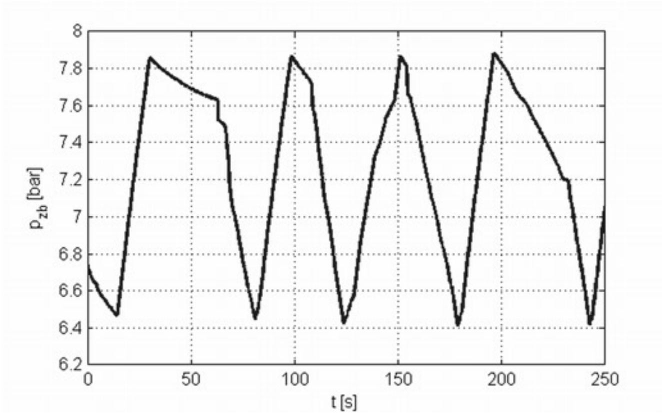


Fig. 6. Example variation of pneumatic system reservoir pressure p_{zb} during testing of the Visteon 51 10 018 unloader valve

4.3. The coupling head pressure test

Testing of the value of pressure p_s at the control coupling head and pressure p_z at the supply coupling head of the dual-line system and pressure p_v at the single-line coupling head involved cyclic braking and releasing with the engine operating. A fragment of the example time variation of the brake pedal force F_p, the hydraulic braking system pressure p_h, the reservoir pressure p_{zb} and the coupling heads pressures, as recorded during the testing of tractor, is shown in Fig. 7.

As determined from the measurements, the pressure variations in the dual-line system were contained in the range of 6.43÷7.87 bar, while in the single-line system, in the range of 6.08÷6.41 bar. The reduction of the minimum controlled pressure value in the dual-line

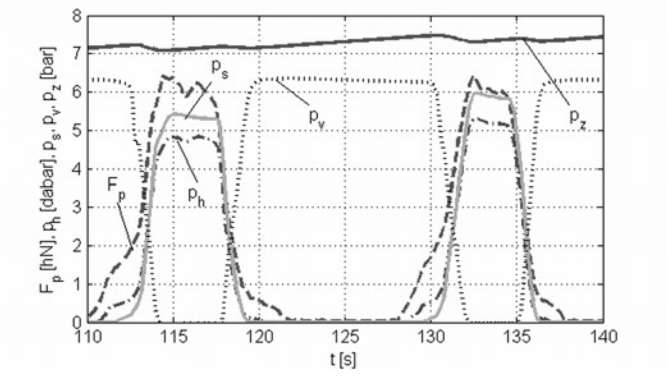


Fig. 7. Example variations of measured quantities during testing the air system coupling head pressure: F_p – brake pedal force; p_s, p_z, p_v – control, supply and single-line coupling head pressure, respectively; p_h – service brake system hydraulic pressure

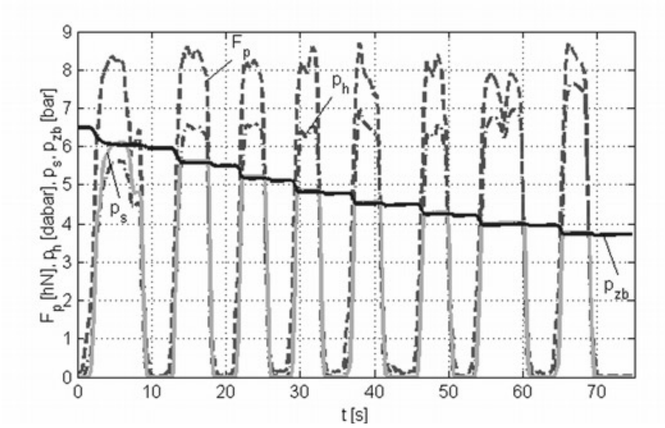


Fig. 9. Example variation of quantities recorded during testing the capacity of the compressed air reservoir in the dual-line system; F_p – brake pedal force, p_h – hydraulic system pressure, p_s – control coupling head pressure, p_{zb} – reservoir pressure

system (below the required 6.5 bar) results from the too low unloader valve switch-on value, not conforming to the specification.

4.4. The compressor capacity test

The test involved the recording of changes in the pressure of compressed air filling the reservoir connected to the supply coupling head of the tractor's air system (Fig. 8).

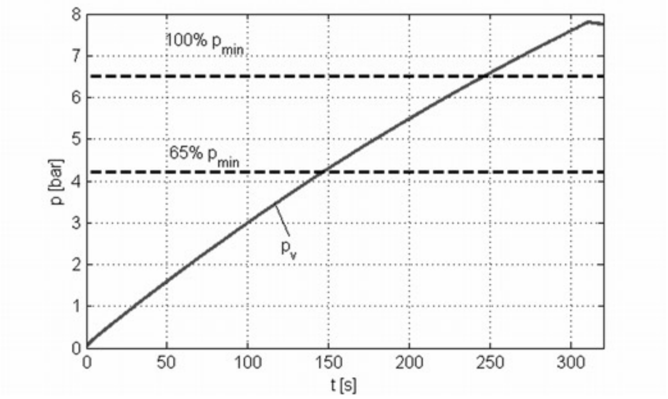


Fig. 8. Example variation of the increase in pressure p_V in the V=60.38 dm³ substitute reservoir during the compressor capacity test.

The time of filling the substitute reservoir with the volume V=60.38 dm³ calculated form relationship (1) was measured from the moment of starting up the heated-up engine up to the point of attaining the required pressures at the maximum engine rotational speed of n_s=2450 rpm. Then, based on 3 measurements, the average time t1 =147.18 s of pressure increasing from zero to 65% of the minimum controlled pressure value and the time t_2=246.23 s needed for attaining 100% of this pressure were determined. The obtained values are less that the maximum values – 360 s and 540 s, respectively – permissible for tractors designed for towing trailers.

4.5. The air reservoir capacity test

The verification of the correctness of selection of the compressed air reservoir in the dual-line system involved carrying out 8 consecutive full brakings with the compressor not operating, and measuring the pressure in the 0.5 dm³ control reservoir connected to the control coupling head (Fig. 9).

From 3 recorded control reservoir pressure measurements, the average pressure values of p_1=6.08 bar after the first braking and p_8=3.81 bar after the eighth braking were determined. The test results indicate the correct selection of the compressed air reservoir pressure.

4.6. The test of the control unit response time

The response time of the tractor's braking system control unit was determined from the recorded variations in the brake pedal force and the pressure at the end of the 2.5 m-long 13 mm-diameter line (simulating the trailer control line) connected to the control coupling head when testing the dual-line system or to the supply & control coupling head when testing the single-line system. Example variations of the measured quantities are shown in Fig. 10.

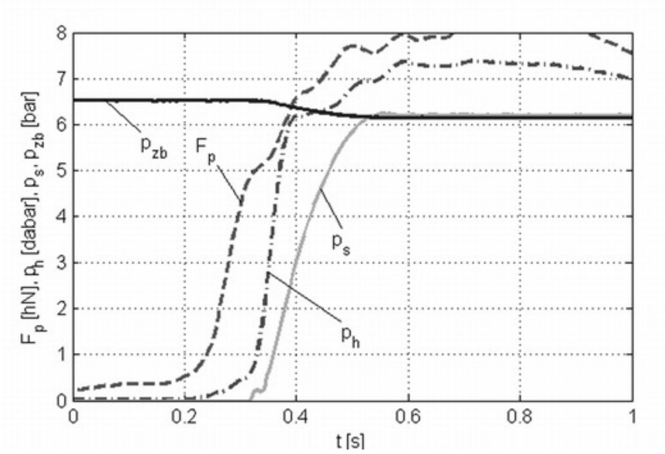


Fig. 10. Example variation of the quantities recorded during testing the tractor response time in the dual-line system: F_p – brake pedal force, p_h – hydraulic system pressure, p_s – control coupling head pressure, p_{zb} – reservoir pressure

Next, based on the recorded variations, the response time t_{10} and t_{75}, i.e. the time of attaining 10% and 75% of the asymptotic pressure value, respectively, was determined as a function of brake pedal application time, starting from the shortest possible applications and then gradually increased up to about 0.4 s. After determining the linear regression equations of the response time t_r as a function of the brake pedal force change time t_f by the least squares method (Fig. 11), the response time corresponding to the time of actuation under sudden braking conditions, i.e. at t_f=0.2 s, was calculated.

The determined response time values t_{10}=0.17 s and t_{75}=0.27 s are lower than the permissible values, which evidences the correct selection of the elements determining the dynamic characteristics of the dual-line braking system of the tractor tested.

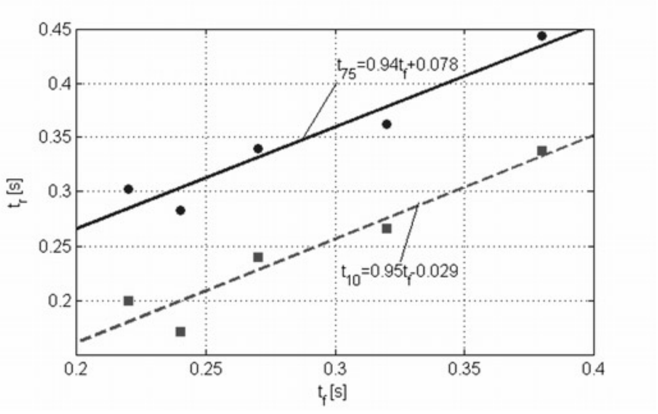


Fig. 11. The effect of the brake pedal force increase time tf on the response time t_{75} (R^2=0.9275) and t_{10} (R^2=0.9122) of the dual-line braking system

4.7. The test of the service braking system efficiency

Example variations of voltage from the fifth wheel pulse sensor, the electronic decelerator and the brake pedal pressure sensor, as recorded during braking of an agricultural tractor using the described measuring system, are shown in Fig. 12.

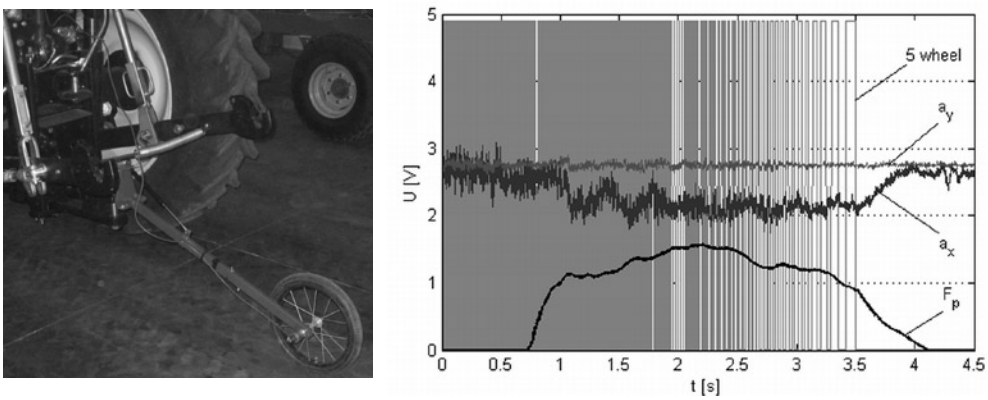


Fig. 12. An example of using the measuring system in braking efficiency tests

Next, based on the recorded quantities, the braking distance, velocity and deceleration were determined as a function of time (Fig. 13) and main braking system efficiency evaluation indicators, such as the stopping distance and the mean fully developed braking deceleration, were calculated. In the performed tests, accelerations of up to 5.22 m/s^2 (as measured with a decelometer), so greater than the minimum value of 4.5 m/s^2 required after the revision of the regulations, were achieved.

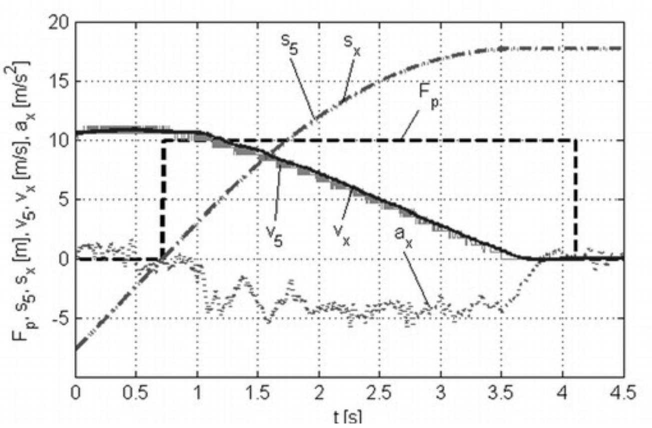


Fig. 13. Example measurement result obtained in the service braking system efficiency test: F_p – brake pedal force indicator; s_5, v_5 – distance and velocity as determined from the fifth wheel measurements; a_x, v_x, s_x – acceleration, velocity and distance, as determined from the electronic decelerator measurements

5. Summary

The agricultural tractor braking system approval testing programme described in the paper covers the most important aspects of diagnosing the brakes of a tractor and its air system for supplying and controlling the brakes of the towed vehicle. The conditions and requirements assumed in individual diagnostic tests are consistent with the proposals of the new Regulation for testing agricultural vehicle brakes [7]. In the authors' view, the proposed provisions should consider the testing of the response time of an agricultural tractor's air system control circuit similarly as is the case in utility vehicles designed for towing trailers.

The presented braking system diagnosing methodology may be either wholly or partially used for developing a programme of diagnostic testing of both newly manufactured agricultural tractors (by qualification testing on production lines) and operated ones (within periodical technical testing on the Vehicle Inspection Stations). This would enable any vehicles with inoperative brakes and air braking systems to be removed from the road traffic, which could improve the safety of road traffic involving tractors coupled with trailers and agricultural machinery.

Literatura

1. Biedrzycki L, Opasewicz W. Powietrzne układy hamulcowe. Budowa i diagnostyka. Poradnik serwisowy. Wydawnictwo Instalator Polski, 2003.
2. Bil H, Kieracińska A, Okniński M. Propozycja zmiany w nadzorowaniu instalacji sprężonego powietrza ciągników rolniczych dla poprawy bezpieczeństwa ruchu drogowego w Polsce. VIII Konferencja hamulcowa '2007 . Hamulce pojazdów drogowych. Łódź 8-9, 11, 2007.
3. Czaban J, Kamiński Z. Badanie skuteczności hamulców ciągników rolniczych. Archiwum Motoryzacji, 2010; 1: 15-25.
4. Dodd M, Bartlett R, Knight I. Provision of information and services on the subject of the performance requirements, testing methods and limit values for braking systems of agricultural and forestry tractors, their trailers and interchangeable towed machinery – final report. TRL Unpublished Project Report No. UPR/VE/064/07, UK, Wokingham, 2007.
5. Dubowski A, Pawłowski T. Medium size road units – an innovative approach for improving efficiency and safety of agricultural transportation in Poland and Europe as well. Increasing work efficiency in agriculture, horticulture and forestry. XXXI CIOSTA-CIGR V Congress Proceedings 2005, pp. 154–161.
6. Directive 2003/37/EC of the European Parliament and of the Council of 26 May 2003 on type-approval of agricultural or forestry tractors, their trailers and interchangeable towed machinery, together with their systems, components and separate technical units and repealing Directive 74/150/EEC.

7. Draft Regulation of the European Parliament and the Council on the braking systems of agricultural or forestry tractors, their trailers and interchangeable towed machinery, amending Directive 2003/37/EC, Council Directive 89/173/EEC and repealing Council Directive 76/432/EEC. (17.11.2008) http://circa.europa.eu/Public/irc/enterprise/automotive/library?l=/agricultural_tractors/meeting_november_2008/99rev16_v171108pdf/_EN_1.0_&a=d.
8. ECE Regulation No. 13. Uniform provisions concerning the approval of vehicles of categories M, N and O with regard to braking. UN Economic Commission for Europe, Geneva, Switzerland, 2001.
9. Kamiński Z, Czaban J. Propozycja programu badań układów hamulcowych ciągników rolniczych (Proposition of exploration program of braking systems of agriculture tractors). MOTROL, 2006; 8: 92–100.
10. Polska norma PN-90/R-36123. Ciągniki. Przyczepy i przyczepiane maszyny rolnicze. Powietrzny jednoprzewodowy układ przenoszący hamulców. Wymagania i badania.
11. Przyczepy leśne – przegląd. Drwal, 2006; 1: 22–25.
12. Radlinski RW, Flick MA. Tractor and trailer brake system compatibility. SAE Transactions, paper no. 861942, 1986.
13. Rozporządzenie Ministra Infrastruktury z dn. 30 grudnia 2003r. w sprawie homologacji ciągników rolniczych (Dz. U. z dnia 15 stycznia 2004 r.). Załącznik nr 3. Wykaz wymagań obowiązujących w homologacji typu. Część B.
14. Scarlett A. In-service assessment of agricultural trailer and trailed appliance braking system condition and performance. The Agricultural Trailer Braking Study. RR697 Research Report, 2009 (http://www.hse.gov.uk/research/rrpdf/rr697.pdf)

Zbigniew KAMIŃSKI, Ph.D., Eng.
Jarosław CZABAN, Ph.D., Eng.
Faculty of Mechanical Engineering
Bialystok University of Technology
ul. Wiejska 45C, 15-351 Białystok, Poland
E-mail: z.kaminski@pb.edu.pl, j.czaban@pb.edu.pl

Article citation info:
TEKCAN A T, KAHRAMANOĞLU G, GÜNDÜZALP M. Incorporating product robustness level in field return rate predictions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 327–332.

A. Tarkan TEKCAN
Gürmen KAHRAMANOĞLU
Mustafa GÜNDÜZALP

INCORPORATING PRODUCT ROBUSTNESS LEVEL IN FIELD RETURN RATE PREDICTIONS

PRZEWIDYWANIE RZECZYWISTEGO WSKAŹNIKA ZWROTÓW TOWARU Z UWZGLĘDNIENIEM POZIOMU ODPORNOŚCI PRODUKTU

Reliability and return rate prediction of products are traditionally achieved by using stress based standards and/or applying accelerated life tests. But frequently, predicted reliability and return rate values by using these methods differ from the field values. The primary reason for this is that products do not only fail due to the stress factors mentioned in the standards and/or used in accelerated life tests. There are additional failure factors, such as ESD, thermal shocks, voltage dips, interruptions and variations, quality factors, etc. These factors should also be considered in some way when predictions are made during the R&D phase. Therefore, a method should be used which considers such factors, thus increasing the accuracy of the reliability and return rate prediction. In this paper, we developed a parameter, which we call Robustness Level Factor, to incorporate such factors, and then we combined this parameter with traditional reliability prediction methods. Specifically, the approach takes into account qualitative reliability tests performed during the R&D stage and combines them with life tests by using Artificial Neural Networks (ANN). As a result, the approach gives more accurate predictions compared with traditional prediction methods. With this prediction model, we believe that analysts can determine the reliability and return rate of their products more accurately.

Keywords: *reliability and return rate estimation, artificial neural networks, defining different failure types, product maturity level, product robustness level, field failures, product level testing, board level testing, design quality.*

Niezawodność i wskaźniki zwrotów towaru przewiduje się tradycyjnie przy użyciu norm obciążeniowych i/lub stosując przyspieszone badania trwałości. Jednakże, często wartości niezawodności i wskaźnika zwrotów przewidywane za pomocą tych metod różnią się od ich wartości rzeczywistych. Główną tego przyczyną jest fakt, że produkty nie ulegają awarii wyłącznie pod wpływem czynników obciążeniowych wymienianych w normach i/lub wykorzystywanych w przyspieszonych badaniach trwałości. Istnieją dodatkowe czynniki wpływające na intensywność uszkodzeń, takie jak wyładowania elektrostatyczne, wstrząsy termiczne, spadki, przerwy w dostawie i zmiany napięcia, czynniki jakościowe, itp. Te czynniki także powinny być w jakiś sposób uwzględnione przy dokonywaniu predykcji na etapie badań i rozwoju (R&D). Dlatego też zwiększenie trafności predykcji niezawodności i wskaźników zwrotów towaru wymaga metody, która uwzględniałaby tego typu czynniki. W niniejszej pracy opracowaliśmy parametr, nazwany przez nas "czynnikiem poziomu odporności", który pozwala na uwzględnienie takich czynników, a następnie wykorzystaliśmy ów parametr w połączeniu z tradycyjnymi metodami przewidywania niezawodności. W szczególności, przedstawione podejście bierze pod uwagę jakościowe badania niezawodnościowe wykonywane na etapie R&D łącząc je z badaniami trwałościowymi przy użyciu sztucznych sieci neuronowych ANN. Dzięki temu, w podejściu tym uzyskuje się bardziej trafne predykcje niż w tradycyjnych metodach prognozowania. Jesteśmy przekonani, że użycie powyższego modelu predykcyjnego umożliwi analitykom bardziej trafne wyznaczanie niezawodności oraz wskaźników zwrotów wytwarzanych przez nich produktów.

Słowa kluczowe: *ocena niezawodności i wskaźnika zwrotów produktu, sztuczne sieci neuronowe, definiowanie różnych typów uszkodzeń, poziom dojrzałości produktu, poziom odporności produktu, awarie w warunkach rzeczywistych, badania na poziomie produktu, badania na poziomie płyty testowej, jakość konstrukcyjna.*

1. Introduction

Consumer electronics reliability has become a major concern for manufacturers in recent years, as consumer electronics prices have dropped quite dramatically. The cost of warranty failure support has become a greater percentage of the profit margin. In Europe, the average cost per failure, including logistics costs, is greater than $150 [13]. Due to this, manufacturers strive to increase the reliability of their products in order to decrease service costs. On the other hand, this improvement causes an increase on design and manufacturing cost.

Companies attempt to reach to the optimum reliability value point, which can only be reached with an accurate reliability and return rate estimation. This estimation should be performed in the R&D stage, prior to mass production. System designers need reliability information during the design phase, to determine the initial, maintenance and total system costs, as well as system-level reliability and availability [5] to perform design changes if needed.

In addition to this, with accurate return rate estimation companies can determine the number of spares to be produced for their services, determine the size of the stock area for spares and also estimate their risk when a product is introduced into market earlier than planned.

Various methods and standards are available for predicting reliability and return rate [4, 6, 8, 10, 12], such as stress based standards which are widely used. Stress based standards, like MIL-HDBK-217F, mainly define the reliability and failure rate of components according to the stress levels on the components. Defined stress factors are temperature, voltage and power dissipation [3]. Additionally, many companies perform accelerated life tests at higher stress and usage rates [15] and analyze the test data with statistical distributions. To accelerate the failure mechanism; temperature, relative humidity, voltage, temperature cycling and vibration are typically used in electronics as the stress factors [1]. The results of these tests are then used to perform reliability predictions under field stress conditions. But frequently, predicted reliability and return rate values by using stress based standards or utilizing accelerated life tests turn out to be significantly different from the actual reliability and return rate value in the field [7]. One reason for this is that sufficient number of samples/prototypes for testing cannot be afforded therefore, accelerated life tests are performed with small sample sizes [11], thus increasing the uncertainty of the predictions. In addition, and in our experience, another significant reason is that the stress factors considered in the standards and used in accelerated life tests are not the only contributors to failures observed in the field during the life period of the product. For example, some additional failure factors for a consumer electronics product can be:

- Electro Static Discharge,
- Inrush Current,
- Voltage Dips-Interruptions-Variations,
- Lightning/Surge Voltages,
- Loose Plugs etc.

During development phase, there is a variety of qualitative tests performed to consider the above mentioned failure factors. However, the outcome of such tests is not considered in the final reliability predictions. Therefore, a method which would consider the outcome of such tests and combine them with the other prediction methods (e.g. accelerated life tests) is needed.

In this paper, a unique approach for determining reliability and field return rate in R&D phase is presented, by introducing a new parameter called Robustness Level Factor (RLF a.k.a. maturity level [14]). This new parameter is obtained by applying a set of electrical, environmental and mechanical tests in R&D phase and prior to mass production. This set of tests typically simulates different stress factors and failure mechanisms faced in the field, which are not life/durability related. These set of tests may also include approval and validation tests at both board and product level. Once the Robustness Level Factor is obtained, Artificial Neural Networks (ANN) is used to combine this parameter with traditional predictions. To demonstrate the methodology and the accuracy of the model, a real life case study on 4 LCD TFT TV projects is given.

2. Determination of the Robustness Level Factor

To determine the Robustness Level Factor (RLF), a set of tests which will simulate the non-life related failure modes which are possible in the field should be created, i.e. robustness tests. These tests should be determined according to type, specification, usage conditions, etc., of the product. In addition, the tests and the corresponding failures found should be assigned numerical values according to their severities in order to determine the risk of the observed failures during testing, relative to the robustness of the final product. These numerical values can be decided by analyzing field returns of similar products.

The calculation method of the RLF for an LCD TFT TV set is given as a real case study, to demonstrate the concept. For this product, the test set consists of electrical, environmental and mechanical tests and these tests can further be grouped as pass/fail tests, early life period tests and design verification tests [9]. All tests have “scoring points” and these points are given according to the importance of the tests. The importance of the tests can be decided by considering the specifications and usage environment, analyzing production line failures, field returns from similar projects etc. As an example, if a product is designed to operate in high temperature environments then a high temperature test will be assigned more scoring points. If a test is thought to be more effective on finding failures, then this test will have higher scoring points.

In addition to this, the failures found during testing are grouped according to their severities, as “showstopper,” “high,” “medium,” and “low”. These failure severities are assigned “losing points” [2]. These points are also decided by analyzing field returns of similar projects, [9]. For example, in our application ESD (electrostatic discharge) is a common failure cause (as observed in our field returns), thus, as we will show later, the ESD test and any failures during that test will be assigned high scoring points and severity.

The scoring points and the losing points affect the result of the prediction method. It is very important to determine the scoring points of the tests and the losing points of the failure severities correctly. If the scoring points and the losing points are determined correctly the results of the prediction method will be closer to the actual value. In our application, the scoring points and the losing points were decided after very long tests and analysis of service data.

In our company, we classified the different robustness tests as follows:

2.1. Pass/Fail Tests

Pass/Fail tests are also referred to as “reliability approval tests”. The main objective of such tests is to find major design flaws. These tests are performed on a product level, and usually with a small sample size. Table 1 provides a list of such tests and the associated scoring points used in our application.

Table 1. List of Pass/Fail Tests

Test Category	Test Name	Scoring Points
Electrical	Voltage Current Stress Test	100
	Temperature Stress Test	100
	Open/Short Circuit Test	100
	ESD Test	100
	Surge Test	25
	Lightning Test	50
	Voltage Dips, Interruption and Variation Test	50
	Power On/Off Test	50
	Inrush Test	75
Environmental	Heat-Run Test	100
	High Temperature Test	50
	Low Temperature Test	50
	High Humidity Life Test	50
Mechanical	Vibration Test	25
	Wall Holder Strength Test	25
	Drop Test	50
	Total	**1000**

2.2. Early Life Period Tests

Early life period (ELP) tests are performed with minimum 20 samples, on a board level. These tests are performed to determine component quality problems, assembly problems, solder-joint problems and failures occurred in early life, for example infant mortality failures. An example of such tests used in our organization is given in Table 2.

Table 2. List of Early Life Period Tests

Test Category	Test Name	Scoring Points
Environmental	Thermal Cycling Test	75
	High Temperature High Humidity Test	50
	Thermal Shock Test	50
Mechanical	Random Vibration Test	50
	Total	**225**

2.3. Design Verification Tests

Design verification tests (DVT) are not the same as the pass/fail tests mentioned above. These tests provide feedback to designers about the weakest points of the design. DVTs are performed with large sample sizes and at a product level, where the test period is longer than pass/fail tests. The main purpose of DVTs is to determine minor design problems. In addition, in these tests combined stress factors are used to accelerate failure mechanism. The list of design verification tests and corresponding scoring points used in our company is given in Table 3.

Table 3. List of Design Verification Tests

Test Category	Test Name	Scoring Points
Electrical	Powered / Unpowered Temperature Cycling Test	100
	ESD Step Stress to Failure Test	50
Environmental	Combined High Temperature High Humidity Test	50
	Thermal Shock Test	75
	Temperature Step Stress to Failure Test	50
	Operational High / Low Temperature Humidity Test	50
	High Humidity Storage Test	25
	Temperature Cycle Test	50
Mechanical	Constructional Inspection Test	50
	Unpackaged Shock Test	50
	Random Vibration Step Stress to Failure Test	25
	Total	**575**

2.4. Total Scoring Points and Losing Points

After the scoring points of the tests are decided, total scoring points are obtained

Then, the "losing points" of the failure severities are determined. As mentioned above, these losing points are decided based on field experience and criticality of the failure based on design specifications.

Table 4. Total Scoring Points

Test Type	Scoring Points
Pass / Fail Tests	1000
Early Life Period Tests	225
Design Verification Tests	575
Total Scoring Points (TSP)	**1800**

Table 5. Losing Points and Failure Severities

Failure Severity	Losing Points
Showstopper (S)	120
High (H)	45
Medium (M)	24
Low (L)	9

When the severities of the failures are decided, total losing points of the project can be calculated from equation (1).

$$TLP = (A \cdot S) + (B \cdot H) + (C \cdot M) + (D \cdot L) \tag{1}$$

TLP: Total Losing Points
A: Number of "Showstopper" Failures
B: Number of "High" Failures
C: Number of "Medium" Failures
D: Number of "Low" Failures

2.5. Calculation of the Robustness Level Factor, RLF

By using equation (2), the parameter, "Robustness Level Factor" (RLF), is calculated.

$$RLF = 1 - \frac{TLP}{TSP} \tag{2}$$

RLF: Robustness Level Factor
TLP: Total Losing Points
TSP: Total Scoring Points

Robustness Level Factor is a number between 0 and 1. This parameter expresses the robustness of the final product related to the failure causes mentioned above. With this parameter, we will attempt to predict the reliability and the return rate of the final product by combining it with life test results and predictions.

3. Combining RLF with life predictions

At this point, the challenge is to combine the computed Robustness Level Factor with predictions made from life tests. As mentioned previously, the motivation is to consider these qualitative failures in our predictions since they represent possible failures in the field. Failing to consider the design robustness in any predictions could result in less accurate estimations. Since the RLF is a qualitative factor (i.e. it does not represent an actual probability) it cannot be easily related to field failures. In order to so, we chose ANN, where a relationship between RLF and life test predictions vs. actual field return rate (based on past projects) can be established/learned. In other words, from past projects, RLF and reliability predictions based on life tests will be the inputs and the actual field return rate the output, and an ANN will be used to "learn" the function between them. Based on the established function, the RLF and the life test reliability prediction of the product under development, we will infer its field return rate.

In engineering and science, ANN are used whenever a function between inputs and outputs needs to be established, where such function is very complex to be determined with other methods or non-applicable (e.g. linear regression). The problem under investigation of combining robustness tests results and life test results falls under this category where no known relationship exists thus using ANN offers an approach to establish it

4. Reliability and return rate estimation

Traditionally, reliability and field return rate predictions were made based on prediction standards (e.g. MIL-217F) and/or life test results. In this paper we consider accelerated life tests where samples are tested at different stress conditions and the results are extrapolated to use level stress conditions. Predicting reliability based on accelerated life tests is a very useful approach; however, such predictions only consider life related failure mechanisms such as electromigration, thermal fatigue, corrosion, etc. In addition, predictions made under this approach are sensitive to the test's sample size. For smaller sample sizes this can result in high uncertainty in the predictions.

In order to increase the accuracy in the predictions and to take into account the robustness level of the design, ANNs will be utilized. In other words, the proposed methodology aims to improve the field return rate prediction by taking into account the accelerated life test results and the RLF. ANN requires existing inputs and outputs to be provided and based on those a function is built. The existing inputs and outputs are called the "training set." In other words, these are the set of values where the algorithm will learn the pattern. In our application, the training set will be the RLF, reliability predictions based on life tests and actual field returns of past projects. The higher the number of past projects used in the training set the more accurate the relationship between inputs and outputs is expected to be. It should also be noted that the proposed methodology and the past projects used apply when making predictions for similar products and for products where history exists (i.e. evolutionary designs).

There are different techniques and algorithms for creating ANNs which are beyond the scope of this paper. In our application and since we only consider two inputs and one output, we used a simple single layer ANN. A real life case study for a LCD TFT TV set is given in the following section.

5. Real life case study

A real life case study to predict reliability and 1st year field return rate values of 3 LCD TFT TV Set projects, with CCFL and LED panels, is given by using 4 older projects' RLF values, accelerated life test predictions, and actual field return rate data (Table VI). These 3 products are currently in the field for more than a year, and their actual 1st year return rates will be provided as a comparison to the predicted ones, to illustrate the applicability of the model.

Table 6. RLF values, accelerated life test results and actual fiedl return rate values

Project	RR_{ALT} (%)	90% 1S UPPER RR_{ALT} (%)	RLF	AFRR (%)
1	4.00	7.98	0,6494	4.19
2	1.82	9.26	0.8275	2.60
3	0.26	4.60	0.9360	0.87
4	2.83	11.75	0.8950	1.57

In Table VI, each project number refers to a real TFT LCD TV product. RRALT denotes the estimated return rate (50% confidence level) calculated by applying accelerated life tests, 90% 1S UPPER RRALT denotes the 90% upper one-sided bound return rate, and RLF denotes "robustness level factor" calculated by applying the tests described in Section II and using the same scoring points provided in that same section. AFRR denotes actual field return rate.

There are a few observations that can be made by examining the information provided in Table VI.

1. The upper bounds are significantly higher than the estimated values. This is due to the small sample sizes used during the ALT.
2. The upper bounds are significantly higher than the actual field return rates.
3. The estimated values (i.e. 50% confidence level) are closer to the actual field return rates.

One could observe that the actual field return rates are below the upper bound computed from the ALT analysis, which is of course the reason of using confidence bounds, i.e. to contain the uncertainty due to sample size. The statement that can be made is that we are 90% confident that the true field return rate will be below the upper bound. However, it can be seen that this bound is very conservative and additionally we need to keep in mind that it does not consider any robustness related failures in the field. In other words, even though the true values given in Table VI are contained within the bounds, one needs to be careful because it is no indication that the robustness related failures are included in the estimation. There may very well be situations where if sufficient number of samples were tested during the ALT, the upper bound would be close to the estimate and could also be below the actual.

We would now like to better understand the relationship between the ALT predictions and the robustness of the product in order to make more intelligent field return rate predictions. For this, we will use ANN as described in Section III. Table VII provides the ALT results and the RLF values for 3 projects currently in the field over 1 year. The actual field return rates are known but we will use the proposed methodology to estimate them and later compare them to the actual values.

Table 7. RLF values and accelerated life test results

Project	RR_{ALT} (%)	90% 1S UPPER RR_{ALT} (%)	RLF
5	0.48	0.66	0.9100
6	0.46	19.55	0.8856
7	1.21	13.73	0.9100

In Table VII we can see that for Project 5 it is expected that the true field return rate will be below 0.66% based on the ALT analysis. In this project it can also be seen that the 90% upper bound is close to the estimated value. This is due to the fact that sufficient samples were tested and for sufficient duration during the development of this product. On the other hand, in projects 6 and 7, less time and samples were available during development and this is reflected in the width between the estimated values and the upper bounds. Based on the history of these products (these 3 products are evolutions of the previous 4 products given in Table VI), we do not expect the actual field return rate to be as high as predicted by 90% upper bound estimates. In addition, our robustness tests have shown that the robustness of these 3 products is high as reflected in the RLF values.

To summarize, we have high confidence in the prediction for project 5 and the actual field return rate is expected to be close to this prediction. However, we need to keep in mind that this prediction does not take into account any robustness related failures which may occur in the field. On the other hand, we are more uncertain about the ALT predictions for projects 6 and 7 and we believe (based on history) that the actual field return rate should be less than these predictions, which also do not include robustness related failures. Finally, all 3 projects have quite high RLF values thus we expect less robustness related failures in the field.

5.1. Analysis and results

To apply the methodology described in section III, we will use the values provided in Table VI as the "training set" for the ANN algorithm. Based on this training set, the ANN will create a relationship between the inputs which are the ALT estimates and the RLF values, and the output which is the AFRR. We will use the 90% upper confidence bound estimate as the ALT input since it provides an estimate which contains the uncertainty of the result.

An ANN was created using the training set described above. We used a single layer ANN with 0 neurons in hidden layers, a minimum weight delta of 0.0001, a learning rate of 0.3, and a zero-based Log-sigmoid-function for the activation function. Based on this ANN and the training set of Table VI, the field return rate of the 3 projects with inputs given Table VII was predicted, and is given in Table VIII.

Table 8. Predicted field return rate values

Predicted Field Return Rate Values (%)	
$PFRR_5$	0.929
$PFRR_6$	2.277
$PFRR_7$	1.459

The following observations can be made from the results of Table VIII;

1. The predicted field return rate for Project 5 is higher than the 90% upper bound value based on ALT analysis. This outcome is actually reasonable, since, as we mentioned previously, the ALT does not consider any robustness related failures where in the field we do expect to see such failures.
2. For projects 6 and 7 the predicted field return rates are much lower than the 90% upper bound value based on ALT analysis. This is also a reasonable outcome as these projects have historically much less failures than what was predicted by the ALT analysis and in addition, they also have a high RLF value which implies consistency and/or improvement over past projects thus similar expectations regarding robustness related failures.

5.2. Comparison of predictions with actual field return rates

As mentioned before, these 3 products are already in the field for more than a year. To illustrate the accuracy of the proposed model and prediction, the actual field returned rates (as obtained from our service department) are given in Table IX and are compared to the predicted values in Table VIII.

As it can be seen from Table IX, the results of the predicted field return rate values by using proposed method are very close to the actual field return rate values and thus more realistic.

Table 9. Actual and predicted 1st year field return rate values

Project	Predicted Field Return Rate (%)	Actual Field Return Rate
5	0.929	1.19
6	2.277	1.14
7	1.459	1.47

6. Conclusions

Reliability and return rate predictions are currently performed mainly by using stress based standards or applying accelerated life tests. However, these methods may not capture every failure reason seen in the field, and specifically robustness related failures. In addition, there is typically a variety of robustness tests performed during development whose outcome indicates the likelihood of observing such failures in the field. Traditionally, the lessons learned and the outcomes of these tests are not taken into consideration when field predictions are performed. It is reasonable to assume that such information regarding a product's robustness should have an influence on the field return rate. For these reasons, we first proposed a parameter to quantify a product's robustness, RLF, using scoring points for the different robustness tests (as described in section II), and we then used this parameter in conjunction with life test results in order to make more realistic field return rate predictions. To do so, we used information from past projects and Artificial Neural Networks in order to create a relationship between the life test results, the RLF and the actual field return rate. The choice of ANNs as a technique was based on the fact that we are not certain about the form of the relationship between these inputs and the output, which could also vary by the application, the product, the historical information, etc. Thus, ANNs provide a general approach which can be used in a variety of products and industries.

The methodology has been used internally in our company and has been proved to be a very useful approach by providing more realistic predictions, which was also demonstrated in this paper by the provided case study. Further utilization of this approach by other industries or reliability engineers would be helpful in order to determine its applicability and value. It should also be noted that our presented approach is an attempt to consider this very real case of including information regarding the robustness of a product and in general any type of qualitative information and test results in the final field return rate prediction. We hope that this approach can also generate interest in this topic and provide stimulation for further research. In fact, the authors are currently also considering other approaches such as the use Bayesian statistics. The research is in its early stages thus no comparisons and results could be reported at the time of this writing.

Acknowledgement

This paper is developed from Ph.D. study of the primary author at Dokuz Eylül University, The Graduate School of Natural and Applied Sciences. The authors would like to thank to Vestel Elektronik Reliability Group for their support.

References

1. Chan H A. Accelerated Stress Testing for Both Hardware and Software. Reliability and Maintainability Symposium 2004.
2. De Visser I.M, Yuan L, Nagappan G. Understanding Failure Severity in New Product Development Processes of Consumer Electronics Products. IEEE International Conference on Management of Innovation and Technology 2006.
3. Defense U.D.o. MIL-HDBK-217F Reliability Prediction of Electronic Equipment 28 Feb 1995.
4. Elerath J G, Pecht, M. IEEE 1413: A Standard for Reliability Predictions. IEEE Transactions on Reliability 2012; 61 (1): 125-129.
5. Goel A, Graves R. J. Electronic System Reliability: Collating Prediction Models. IEEE Transactions on Device and Materials Reliability 2006; 6 (2): 258–265.
6. Jackson D S, Pant H, Tortorella M. Improved Reliability-Prediction and Field-Reliability-Data Analysis for Field-Replaceable Units. IEEE Transactions on Reliability 2002; 51 (1): 8–16.
7. Jones J, Hayes J. A Comparison of Electronic-Reliability Prediction Models. IEEE Transactions on Reliability 1999; 48 (2): 127-134.

8. Jones J, Hayes J. Estimation of System Reliability Using a "Non-Constant Failure Rate" Model. IEEE Transactions on Reliability 2001; 50 (3): 286–288.
9. Kahramanoğlu G. Field Return Rate Estimation in R&D Stage with an Indicator. Accelerated Stress Testing Reliability Symposium 2010.
10. Lin D L, Welsher T L. Prediction of Product Failure Rate due to Event-Related Failure Mechanisms. Reliability and Maintainability Symposium 1998.
11. Ma H, Meeker W.Q. Strategy for Planning Accelerated Life Tests With Small Sample Sizes. IEEE Transactions on Reliability 2010; 59 (4): 610–619.
12. Ormon S W, Cassady C R, Greenwood A.G. Reliability Prediction Models to Support Conceptual Design. IEEE Transactions on Reliability 2002; 51 (2): 151–157.
13. Shaw M. LCD TV Reliability Testing: An Effective Approach. Applied Reliability Symposium Europe 2010.
14. Tekcan T, Kirisken B. Reliability test procedures for achieving highly robust electronic products. Reliability and Maintainability Symposium 2010.
15. Yang G. Accelerated Life Tests at Higher Usage Rates. IEEE Transactions on Reliability 2005; 54 (1): 53–57.

A. Tarkan TEKCAN, M.Sc., EEE
Gürmen KAHRAMANOĞLU, B.Sc., ECE
Vestel Electronics AS
Organize Sanayi Bolgesi, 45030, Manisa, Turkey
E-mail: tarkan.tekcan@vestel.com.tr, gurmen.kahramanoglu@vestel.com.tr

Mustafa GÜNDÜZALP, Ph.D., EEE
Department of Electrical-Electronic Engineering
Yaşar University
Üniversite Caddesi, No:35-37, Ağaçlı Yol, 35100, Bornova, İzmir, Turkey
E-mail: mustafa.gunduzalp@yasar.edu.tr

Article citation info:
BUGARIC U, TANASIJEVIC M, POLOVINA D, IGNJATOVIC D, JOVANCIC P. Lost production costs of the overburden excavation system caused by rubber belt failure. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 333–341.

Ugljesa BUGARIC
Milos TANASIJEVIC
Dragan POLOVINA
Dragan IGNJATOVIC
Predrag JOVANCIC

LOST PRODUCTION COSTS OF THE OVERBURDEN EXCAVATION SYSTEM CAUSED BY RUBBER BELT FAILURE

KOSZTY UTRACONEJ PRODUKCJI WYWOŁANE USZKODZENIEM GUMOWEJ TAŚMY TRANSPORTOWEJ W UKŁADZIE MASZYNOWYM DO ZDEJMOWANIA NADKŁADU

In this paper the average malfunction costs (lost production) of the overburden excavation system on the Tamnava - East field open-pit mine caused by the failure of belt conveyor rubber belts which work on the bucket wheel excavator, belt wagon and spreader are determined, i.e. the unit cost of system malfunction per hour of belt conveyor work during belt lifetime. The basis for the calculation of malfunction costs is presented by the proposed methodology for the analysis of rubber belt working time to failure based on the fact that working time until sudden failure (tear, breakthrough) can be described by exponential distribution while working time until gradual failure can be described by normal distribution. The proposed methodology as well as the expression for malfunction cost determination can be used, with appropriate adoptions, in the analysis of the functioning of other open-pit mines for better planning of malfunctions, requirements for spare rubber belts as well reductions in working costs, i.e. they can indicate better (optimal) maintenance strategy.

***Keywords**: Malfunction costs, Rubber belt failure, Failure time distribution.*

W prezentowanym artykule wyznaczono średnie koszty awarii (utraconej produkcji) układu maszynowego do zdejmowania nadkładu, wykorzystywanego w kopalni odkrywkowej Tamnava - East field, spowodowanej uszkodzeniem przenośnikowych taśm gumowych zastosowanych w koparce wielonaczyniowej, samobieżnym przenośniku taśmowym oraz zwałowarce. Koszt zdefiniowano jako jednostkowy koszt awarii układu na godzinę pracy przenośnika taśmowego podczas cyklu życia taśmy. Podstawę obliczeń kosztów awarii stanowiła zaproponowana metoda analizy czasu pracy taśmy gumowej do uszkodzenia, oparta na fakcie, iż czas pracy do nagłego uszkodzenia (rozerwanie, przebicie) można opisać za pomocą rozkładu wykładniczego, natomiast czas pracy do stopniowego uszkodzenia – za pomocą rozkładu normalnego. Proponowana metodologia, jak również równania do wyznaczania kosztów awarii mogą być wykorzystywane, przy odpowiedniej adaptacji, do analizy funkcjonowania innych kopalni odkrywkowych służąc lepszemu planowaniu awarii, zapotrzebowania na zapasowe pasy gumowe oraz redukcji kosztów pracy. Mogą one, innymi słowy, wskazywać (optymalną) strategię utrzymania ruchu.

***Słowa kluczowe**: Koszty awarii, uszkodzenie taśmy gumowej, rozkład czasu uszkodzenia.*

1. Introduction

Within the public company - Electric Power Industry of Serbia (EPS), during the observed period (1991-2009), coal (lignite) was excavated from six open-pit mines in two basic basins, Kolubara and Kostolac. Continuous mechanisation was used for the excavation of lignite and overburden on all of the open-pit mines. The work of the Tamnava – East field open-pit mine located in the Kolubara basin, with an average annual lignite production of 7.5 million tons and excavation of 10 million m^3lm (cubic meters of loose material) of overburden, is observed for the gathering and analysis of data about the work of mechanisation in the sense of malfunctions caused by failures of the rubber belts on the belt conveyors. It is important to underline that almost all the lignite production from the Tamnava – East field open-pit mine was used for the production of electric energy.

The excavation, removal and disposal of overburden are prerequisites for lignite exploitation. The composition of overburden which has to be excavated is heterogeneous and consists of gravel and several types of clay. From 1979 continuous mechanisation was used for overburden excavation on the Tamnava – East field open-pit mine. Mechanisation consisted of several machines such as: a bucket wheel excavator (theoretical capacity Q=4100 m^3lm/h), mobile transfer conveyors (belt wagons), spreaders and a belt conveyor system. The rubber belts with iron cord (St=1600 N/cm), which are used on the conveyors on the machines are 1600 mm wide. While receiving excavated material, in addition to the expected wearout, sudden rubber belt failures occur due to the impact caused by large excavated pieces. Transported material also has a dominant influence on both the reliability and durability of rubber belts (of the same quality). Namely, the lifetime of rubber belts is from 4 to 5 times longer during the transportation of lignite than in the case of overburden transportation [13].

Due to those facts and the subsequent the lack of data for analysing rubber belt failure during lignite transportation in the observed period, in this paper only rubber belt failure on belt conveyors in the case of overburden transportation, which arise more often than in the case of lignite transportation, will be analysed.

In this paper the methodology for the analysis of the rubber belts working time to failure, based on the fact that the working time until sudden failure (tear, breakthrough) can be described by exponential distribution while that until gradual failure can be described by normal distribution, is proposed and verified. Using the proposed methodology the data related to the failure of belt conveyor rubber belts, which work in the overburden excavation system (bucket wheel excavator, belt wagon, spreader), are analysed (statistically processed) for the period from 1991 to 2009.

Six rubber belts of different lengths which work on specified machines are analysed in order to determine the average malfunction costs (lost production) of the overburden excavation system on the Tamnava - East field open-pit mine caused by the failure of belt conveyor rubber belts, i.e. the unit cost of system malfunction per hour of belt conveyor work during belt lifetime.

The overburden excavation systems work 24 hours a day 7 days a week i.e. round the clock, and maximum production can only be achieved through the maximum usage of equipment. However, poor design solutions and other unexpected problems limit their performance and effectiveness. Those problems, caused by inadequate reliability, maintainability characteristics and poor maintenance strategy, lead to unexpected breakdowns and failures which result in huge economic losses [5, 9].

1.1. Literature review

Mining equipment complexity and size are continually increasing and therefore unplanned failures of mining equipment cause higher repair (replacement) costs than the planned maintenance or repair. On the other side, lost production costs are even more important. Those facts underline the importance of a reliability study into the operation of mining equipment [2].

With the beginning of the development of systems sciences, practically after World War II, reliability engineering as one of the main concepts of technical systems assessment experienced the most progressive development. Barabady and Kumar [2] performed reliability analysis for each subsystem of two analysed chrushing plants by using failure data in order to estimate the parameters of theoretical distributions which provide the best fit for characterizing the failure pattern. Uzgoren and Elevli [16] showed that the times between successive failures for the mechanical systems of a dragline are not independent and identically distributed and use the nonhomogeneous Poisson process to describe the trend; i.e. the time between failures as a function of time, in order to predict the time to the next failure and thus determine the expected number of failures and reliability for different time periods. Hoseine et al. [7] analysed the reliability of the water system which plays a critical role in the work of longwall shearer machines. The water system is modelled as a system of three serially connected subsystems. The empirical data about the functioning of the water system are statistically analysed in order to determine which subsystem has the highest reliability importance. In that sense, Uzgoren et al. [15] claim that the reliability function is the base of reliability investigations and hence perform a comparative analysis of twenty studies of mining machinery and equipment reliability, which emphasizes the importance of reliability research especially in the mining industry. There has been no analysis of conveyor rubber belts.

Reliability research and failure analyses of the belt conveyors are very important, due to the costs of the technological processes in which belt conveyors are an integral element. Belt puncture resistance, slit resistance, belt fatigue testing, and the investigation of belt splices are the basic experimental methods of assessment of reliability and remaining rubber belt capabilities – resources [6]. Bindzar et al. [3] propose several numerical methods in order to assess the remaining capabilities of rubber belts due to service quality.

Chookah et al. [4], used the superposition of the probability functions to compose the degradation effects (phenomena: fatigue, corrosion) on oil pipelines. The super-position of the probability functions model unites various phenomena that have the same effect. The same applies in the case of different categories of rubber belt failure and degradation.

The influence factors on the maintenance costs are analysed and investigated by Yue et al. [17], in order to establish reasonable maintenance strategies, improve the efficiency of equipment operation, reduce costs and optimize the design of mining machines. Based on the time-to-failure density function, Huang et al. [8] analysed the influence of the maintainability and the maintenance policy in general on service life. Liu et al. [10] used the Weibull probability density function for reliability simulation so as to optimize the design and reduce the maintenance costs of chain conveyors. Elevli and Demirci [5] analysed the impact of corrective, preventive and mixed maintenance activities (strategies) of mining equipment on reliability and profitability, defining the relationship between reliability and profitability. The common conclusion reached in these articles is a well-known fact – increasing reliability (up to a certain level) can reduce maintenance costs. For rubber belts, it is significant to observe this fact in relation to belt length; i.e. to find the mathematical dependence between maintenance costs, failure rates and belt length.

2. Methodology for failure analysis and indexes of excavation system functioning

From the reliability point of view an overburden excavation system can be considered as a serial system, meaning that the failure of any machine or sub-system leads to the malfunction of the whole system.

Data pertaining to the malfunction of the overburden excavation system caused by rubber belt failure are taken from the maintenance diary (plant records) of the Tamnava – East field open-pit mine for a 19 year period (from 1991 to 2009). Six closed rubber belts of different lengths, which work on the bucket wheel excavator, belt wagon and spreader, are taken into consideration. The lengths of those rubber belts are: L_1=27.6 m, L_2=51.6 m, L_3=75 m, L_4=81.5 m, L_5=106.1 m and L_6=158.2 m.

The basis of the proposed methodology is the idea that the causes of rubber belt failure can be divided into two categories i.e. that the distributions of working time to failure caused by different failure categories are not the same.

Depending on the nature of the cause which leads to the malfunction of the overburden excavation system, the data about rubber belt failure are divided into two categories. The first category “gradual failures” comprises failure caused by:

- worn out belt,
- belt runout,
- belt replacement due to annual maintenance, and
- belt replacement due to capital maintenance,

while the second category “sudden failures” comprise failure caused by:

- damage to junctions,
- breakthrough,
- impact, and
- belt tearing.

Since the data from the maintenance diary only tells us when the failure of each rubber belt occurred (date) and which cause lead to such failure, additional computation to determine the time each rubber belt worked until failure has to be carried out.

In order to obtain the “effective working time per year [h]” for each analysed rubber belt regardless of the nature of failure (Tables 2, 3, 4, 5, 6, 7 and 10), the number of days the rubber belt functioned properly in a specific year is multiplied by the appropriate annual coefficient of time utilisation k_t (Table 1) of the overburden excavation system. The coefficient of time utilisation k_t is calculated based on 8760 working hours per year.

In addition to the values for k_t, Table 1 also shows: the values for the coefficient of capacity utilization k_q, the annual quantities of excavated overburden Q_{ob} [m3lm/year] as well as the annual quantities of excavated lignite Q_l [t/year] on the Tamnava – East field open-pit mine in the observed period (from 1991. to 2009.), which are necessary for the calculation of the average malfunction costs of the overburden excavation system.

2.1. Analysis of rubber belt working time to "gradual failures"

The effective working time per year (obtained as described in the previous chapter), the total working time between failures (interventions) and the mean time to failure caused by gradual failures – $MTTF_{gf}^{L_i}$, for each of the analysed rubber belts, are shown in Tables 2, 3, 4, 5, 6 and 7.

In the same time period the shorter rubber belts receive material more often and perform more deflections around the pulleys then their longer counterparts. In the case of gradual failures those facts lead to faster wearout and therefore to shorter working (mean) time to failure i.e. shorter lifetime, which is shown in Table 8.

Table 1. Indexes of overburden excavation system functioning [1]

Year	Coefficient of time utilisation (k_t)	Coefficient of capacity utilisation (k_q)	Quantity of excavated overburden *Qob* [m³lm/year]	Quantity of excavated lignite *Ql* [t/year]
1991	0.53	0.55	10501264.80	10133116.00
1992	0.47	0.51	9730550.70	11709953.00
1993	0.56	0.56	11381546.80	10650709.00
1994	0.53	0.48	9019914.80	11549667.00
1995	0.48	0.46	8016591.70	11911686.00
1996	0.54	0.41	8875252.10	8510809.00
1997	0.52	0.44	9319011.00	9160856.00
1998	0.38	0.37	7524601.50	8115122.00
1999	0.37	0.44	7945520.70	8127768.00
2000	0.32	0.39	6676672.60	8014754.00
2001	0.33	0.43	7856667.00	5616794.00
2002	0.31	0.53	9574307.60	3239635.00
2003	0.43	0.56	10237852.30	5870371.00
2004	0.43	0.45	7911295.60	4798317.00
2005	0.41	0.44	10213457.80	4859044.00
2006	0.59	0.43	10196516.20	4805146.00
2007	0.58	0.49	10086417.90	5036367.00
2008	0.61	0.44	10397788.70	5029341.00
2009	0.47	0.43	3318868.80	3479869.00
	Average: $\bar{k}_t = 0.466$	Average: $\bar{k}_q = 0.464$	Total: $Q_{ob}^{tot} = 168784098.6$	Total: $Q_l^{tot} = 140619324$

The analysed rubber belts have different lengths and in order to obtain a homogenous (representative) sample for further analysis, it is necessary to recalculate the total working time of the shorter rubber belts according to the longest rubber belt (L_6=158.2 m). Therefore, the mean time to failure ($MTTF_{gf}^{L_i}$) caused by gradual failures, as a function of the rubber belt length is determined (L_i) (Table 8). For that purpose the linear regression model is assumed.

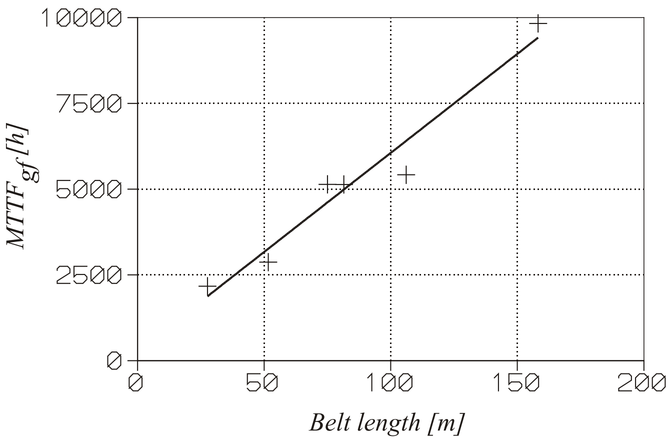


Fig. 1. Mean time to gradual failures as a function of rubber belt length

The parameters of linear regression and the correlation coefficient are determined using statistical software SPSS [14].

The required linear regression, shown on Figure 1, has the following form:

$$MTTF_{gf} = 281.41+57.734\cdot L. \qquad (1)$$

The correlation coefficient, the measure of the strength of linear dependence, has the following value r = 0.976.

The value of the correlation coefficient is very close to one, which means that the dependence between the mean time to failure, caused by gradual failures – $MTTF_{gf}^{L_i}$ and rubber belt length – L_i is absolute (strong). [11]

In order to identify whether the recalculation of the total working time of the shorter rubber belts according to the longest rubber belt makes sense, the existence of statistical dependence between the mean time to failure, caused by gradual failures and rubber belt length also needs to be checked.

Verifying the statistical dependence between the mean time to failure and rubber belt length, for a small sample case ($n<30$), is done by using Student's t-test. Statistical dependence exists if [11]:

$$|t_0|>t_{\alpha/2,n-2}, \qquad (2)$$

where:

$t_0 = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ = 9.02169 – decision statistics (r = 0.976 – correlation coefficient, n= 6),

$t_{\alpha/2,n-2}$ – theoretical value of Student's t distribution (for α= 0.05), i.e. $t_{\alpha/2,n-2}=t_{0.025,4}$= 2.776.

According to the previous expressions, it can be concluded that statistical dependence between the mean time to failure, caused by gradual failures – $MTTF_{gf}^{L_i}$ and rubber belt length – L_i exists because

[1] Source: Public company Electric Power Industry of Serbia.

Table 2. Rubber belt working time to gradual failure – L_1= 27.6 m

No.	Year	No. of days	Effective working time per year [h]	Total working time [h]
1	1992	37	417.36	2258.64
	1993	137	1841.28	
2	1994	242	3078.24	3078.24
3	1994	53	674.16	2010.48
	1995	116	1336.32	
4	1997	114	1422.72	1422.72
5	1998	122	1112.64	1112.64
6	2007	96	1336.32	2361.12
	2008	70	1024.80	
7	2008	88	1288.32	1288.32
8	2008	113	1654.32	1654.32
9	2008	95	1390.80	4323.60
	2009	260	2932.80	
$MTTF_{gf}^{L_1}$ [h]				**2167.79**

Table 3. Rubber belt working time to gradual failure – L_2= 51.6 m

No.	Year	No. of days	Effective working time per year [h]	Total working time [h]
1	1992	135	1522.80	3095.28
	1993	117	1572.48	
2	1993	248	3333.12	4134.48
	1994	63	801.36	
3	1995	238	2741.76	6111.36
	1996	260	3369.60	
4	1996	51	660.96	660.96
5	1996	55	712.80	3208.80
	1997	200	2496.00	
6	1998	107	975.84	1482.00
	1999	57	506.16	
7	2008	254	3718.56	3718.56
8	2009	48	541.44	541.44
$MTTF_{gf}^{L_2}$ [h]				**2869.11**

Table 4. Rubber belt working time to gradual failure – L_3= 75 m

No.	Year	No. of days	Effective working time per year [h]	Total working time [h]
1	1993	149	2002.56	4304.88
	1994	181	2302.32	
2	1995	101	1163.52	1163.52
3	1995	72	829.44	9254.40
	1996	366	4743.36	
	1997	295	3681.60	
4	1997	70	873.60	3618.72
	1998	301	2745.12	
5	1999	143	1269.84	8377.68
	2000	366	2810.88	
	2001	365	2890.80	
	2002	189	1406.16	
6	2007	67	932.64	5310.00
	2008	299	4377.36	
7	2008	67	980.88	3924.96
	2009	261	2944.08	
$MTTF_{gf}^{L_3}$ [h]				**5136.31**

Table 5. Rubber belt working time to gradual failure – L_4= 81.5 m

No.	Year	No. of days	Effective working time per year [h]	Total working time [h]
1	1991	155	1971.60	10038.00
	1992	366	4128.48	
	1993	293	3937.92	
2	1993	72	967.68	4440.24
	1994	273	3472.56	
3	1997	134	1672.32	3186.24
	1998	166	1513.92	
4	1998	197	1796.64	1796.64
5	1999	143	1269.84	7670.88
	2000	366	2810.88	
	2001	365	2890.80	
	2002	94	699.36	
6	2007	35	487.20	4527.84
	2008	276	4040.64	
7	2008	90	1317.60	
	2009	263	2966.64	4284.24
$MTTF_{gf}^{L_4}$ [h]				**5134.87**

the calculated value of decision statistics is greater than the theoretical value of Student's *t* distribution.

In order to obtain a representative sample, the correction (recalculation) of the working times to the gradual failures (TTF) of the rubber belts is done in the following way: each working time to gradual failure of the shorter rubber belts (L_1, L_2, L_3, L_4, L_5) is multiplied by the quotient of mean times to gradual failure, obtained using the linear dependence (1), of the longest (L_6) and the given rubber belt. The obtained results are shown in Table 9.

The verification of the assumption that the corrected sample, shown in Table 9, can be represented by normal distribution:

$$f(t) = \frac{1}{\sigma \cdot \sqrt{2 \cdot \pi}} \cdot \exp\left[-\frac{(t-\bar{t})^2}{2 \cdot \sigma^2} \right] \tag{3}$$

Table 6. Rubber belt working time to gradual failure - L_5= 106.1 m

No.	Year	No. of days	Effective working time per year [h]	Total working time [h]
1	1992	284	3203.52	5542.08
	1993	174	2338.56	
2	1994	285	3625.20	3625.20
3	1994	7	89.04	1183.44
	1995	95	1094.40	
4	1995	270	3110.40	5793.12
	1996	207	2682.72	
5	1996	159	2060.64	5055.84
	1997	240	2995.20	
6	1997	125	1560.00	12421.92
	1998	365	3328.80	
	1999	365	3241.20	
	2000	366	2810.88	
	2001	187	1481.04	
7	2008	94	1376.16	4297.68
	2009	259	2921.52	
			$MTTF_{gf}^{L_5}$ [h]	**5417.04**

Table 7. Rubber belt working time to gradual failure - L_6= 158.2 m

No.	Year	No. of days	Effective working time per year [h]	Total working time [h]
1	1991	324	4121.28	10709.28
	1992	366	4128.48	
	1993	183	2459.52	
2	1993	182	2446.08	5130.00
	1994	211	2683.92	
3	1994	154	1958.88	10077.60
	1995	365	4204.80	
	1996	302	3913.92	
4	1997	6	74.88	14035.44
	1998	365	3328.80	
	1999	365	3241.20	
	2000	366	2810.88	
	2001	365	2890.80	
	2002	227	1688.88	
5	2003	288	2972.16	16748.16
	2004	366	3777.12	
	2005	365	3591.60	
	2006	365	5168.40	
	2007	89	1238.88	
6	2007	276	3841.92	5862.24
	2008	138	2020.32	
7	2008	228	3337.92	6248.16
	2009	258	2910.24	
			$MTTF_{gf}^{L_6}$ [h]	**9830.13**

Table 8. Rubber belt lifetime due to gradual failures

i	Belt length L_i [m]	$MTTF_{gf}^{L_i}$ [h]
1	27.6	2167.79
2	51.6	2869.11
3	75	5136.31
4	81.5	5134.87
5	106.1	5417.04
6	158.2	9830.13

where $\overline{t}$ and σ represent the arithmetic mean and standard deviation of the sample, is done by applying the χ^2 – test, using software specially designed for analysing stochastic variables in transport systems [12]. The cumulative value of the test statistic is χ^2_{ts} = 13.404, obtained by dividing the sample shown in Table 9 into 8 classes. The critical value, i.e. the table value for chi-square distribution for 5 degrees of freedom and significance level α = 0.01 is χ^2_{cr} = 15.086. Therefore the corrected sample (Table 9) can be represented (described) by normal distribution, with the arithmetic mean $\overline{t}$ = 9556.69 and standard deviation σ = 4967.05. (Figure 2)

Verification that the corrected sample given in Table 9 fits normal distribution is also done by the one sample Kolmogorov-Smirnov test, where the p-value of the test is 0.464 and the significance level is 0.05, using SPSS statistical software [14].

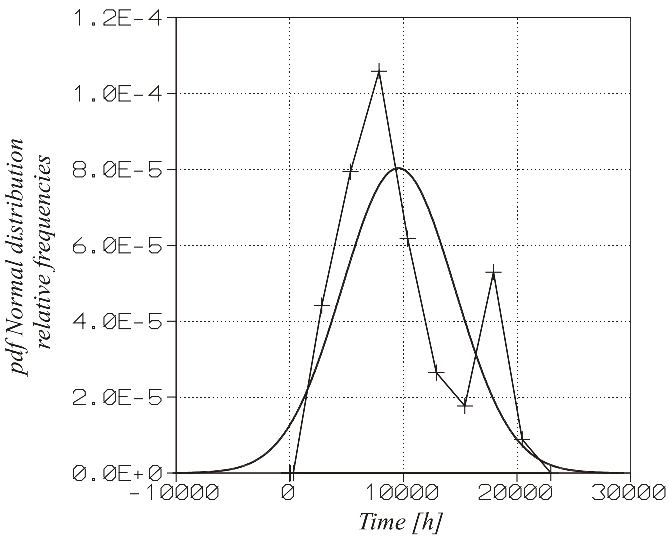


Fig. 2. Pdf of rubber belt working times to gradual failures – normal distribution

2.2. Analysis of rubber belt working time to „sudden failures"

The causes of failure in the second category "sudden failures" are breakthrough, impact, tearing and other damage which in the majority of cases are the result of the stroke exerted by large excavated pieces on the rubber belt. Sudden failures can occur at any moment during the operating of the belt conveyor, regardless of rubber belt length and the time the rubber belt has been in use (new, regenerated etc.). Due to the nature of sudden failures, there is no dependence between the

Table 9. Corrected working times to gradual failures

L_i		TTF [h]	L_i		TTF [h]	L_i		TTF [h]
L_1 = 27.6 m	1	11342.09	L_2 = 51.6 m	1	8937.88	L_3 = 75 m	1	8789.00
	2	15457.82		2	11938.66		2	2375.49
	3	10095.91		3	17647.07		3	18894.12
	4	7144.39		4	1908.58		4	7388.11
	5	5587.28		5	9265.68		5	17104.18
	6	11856.70		6	4279.40		6	10841.09
	7	6469.48		7	10737.66		7	8013.34
	8	8307.41		8	1563.45			
	9	21711.58						
L_4 = 81.5 m	1	18951.70	L_5 = 106.1 m	1	8143.97	L_6 = 158.6 m	1	10709.28
	2	8383.15		2	5327.15		2	5130.00
	3	6015.61		3	1739.04		3	10077.60
	4	3392.05		4	8512.86		4	14035.44
	5	14482.59		5	7429.45		5	16748.16
	6	8548.54		6	18253.74		6	5862.24
	7	8088.63		7	6315.35		7	6248.16

working time and length of the rubber belts and therefore there is no need for sample correction (recalculation).

The effective working times per year of rubber belt due to "sudden failures", as in the previous case, are obtained by multiplying the number of days the rubber belt functioned properly in a specific year by the appropriate annual coefficient of time utilisation K_t (Table 1) of the overburden excavation system.

The effective working time per year for each rubber belt is shown in Table 10, where the column "Total working time [h]" represents the sample which will be tested for consistence with exponential distribution. The mean time to sudden failures $MTTF_{sf}^{L_i}$ is shown in the far right column of Table 10.

The verification of the assumption that the sample, shown in Table 10 (sixth column), can be represented by exponential distribution:

$$f(t) = \lambda \cdot \exp[-\lambda \cdot t] \tag{4}$$

where λ represents the parameter of exponential distribution ("sudden failure" intensity), is done by applying of the χ^2 – test, using also software specially designed for analysing stochastic variables in transport systems [12].

The cumulative value of the test statistic is χ_{ts}^2 = 1.881, obtained by dividing the sample shown in Table 10 (sixth column) into 7 classes. The critical value, i.e. table value for chi-square distribution for 5 degrees of freedom and significance level α = 0.01 is χ_{cr}^2 = 15.086. Therefore the sample shown in Table 10 can be represented (described) by exponential distribution, with the parameter λ = 0.0005. (Figure 3).

The verification that the sample given in Table 10 fits exponential distribution is also done by the one sample Kolmogorov-Smirnov test, where the p-value of the test is 0.888 and the significance level is 0.05, using SPSS statistical software [14].

3. Malfunction costs of the overburden excavation system caused by belt conveyor rubber belt failure

As shown in previous chapters, rubber belt failures which causes malfunctions in the belt conveyor i.e. the functioning of the overburden excavation system, can be divided into two categories: sudden failures (exponentially distributed working times to failure) and gradual failures (normally distributed working times to failure). In both cases it is assumed that such failures are instantaneous while failure clear up is carried out by belt replacement. In the case of gradual failures belt

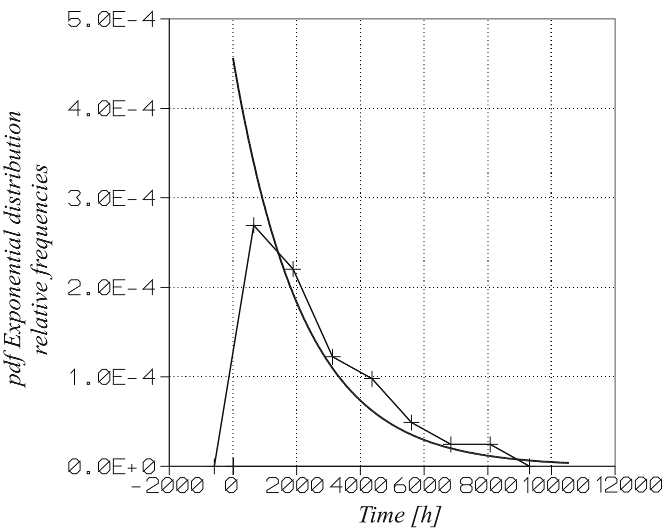


Fig. 3. Pdf of rubber belt working times to sudden failures – exponential distribution

replacements can be planned in advance (smaller replacement costs) which cannot be done in the case of sudden failures. Furthermore, when sudden failures occur the rubber belt resource is not fully used.

The average malfunction costs (lost production) of the overburden excavation system caused by belt conveyor rubber belt failure, i.e. the unit cost of system malfunction per hour of belt conveyor work during belt lifetime can be calculated using the following expression [1]:

$$\bar{C}_{m_i} = \frac{c_{pm} + (c_{npm} + c_{nbr}^{L_i}) \cdot H(MTTF_{gf}^{L_i})}{MTTF_{gf}^{L_i}} \tag{5}$$

where:

Table 10. Rubber belt working times to sudden failure

L_i [m]	No.	Year	No. of days	Effective working time per year [h]	Total working time [h]	$MTTF_{sf}^{L_i}$ [h]
L_1 = 27.6 m	1	1993	122	1639.68	1639.68	**1918.08**
	2	1993	106	1424.64		
		1994	70	890.40	2315.04	
	3	1995	83	956.16	956.16	
	4	1997	168	2096.64		
		1998	103	939.36	3036.00	
	5	1998	51	465.12	465.12	
	6	1998	37	337.44	337.44	
	7	1998	52	474.24		
		1999	87	772.56	1246.80	
	8	1999	32	284.16	284.16	
	9	1999	94	834.72	834.72	
	10	1999	145	1287.60	1287.60	
	11	1999	7	62.16		
		2000	366	2810.88		
		2001	365	2890.80		
		2002	365	2715.60		
		2003	21	216.72	8696.16	
L_2 = 51.6 m	1	1994	302	3841.44		3192.24
		1995	127	1463.04	5304.48	
	2	1997	165	2059.20		
		1998	211	1924.32	3983.52	
	3	1998	47	428.64	428.64	
	4	2008	96	1405.44		
		2009	146	1646.88	3052.32	
L_3 = 75 m	1	1992	216	2436.48		**1913.04**
		1993	121	1626.24	4062.72	
	2	1993	95	1276.80	1276.80	
	3	1994	184	2340.48		
		1995	192	2211.84	4552.32	
	4	1998	64	583.68		
		1999	222	1971.36	2555.04	
	5	2002	176	1309.44		
		2003	59	608.88	1918.32	
	6	2003	41	423.12	423.12	
	7	2003	23	237.36	237.36	
	8	2003	3	30.96	30.96	

Table 10. Rubber belt working times to sudden failure. (continue)

L_i [m]	No.	Year	No. of days	Effective working time per year [h]	Total working time [h]	$MTTF_{sf}^{L_i}$ [h]
L_4 = 81.5 m	1	1998	2	18.24		**2622.88**
		1999	222	1971.36	1989.60	
	2	2002	271	2016.24		
		2003	212	2187.84	4204.08	
	3	2009	102	1150.56		
		2010	95	524.40	1674.96	
L_5 = 106.1 m	1	1993	191	2567.04		**3883.32**
		1994	19	241.68	2808.72	
	2	1994	54	686.88	686.88	
	3	2001	178	1409.76		
		2002	365	2715.60		
		2003	250	2580.00	6705.36	
	4	2007	97	1350.24		
		2008	272	3982.08	5332.32	
L_6 = 158.2 m	1	1996	64	829.44		**2337.04**
		1997	132	1647.36	2476.80	
	2	1997	227	2832.96	2832.96	
	3	2002	138	1026.72		
		2003	77	794.64	1821.36	

$\bar{C}_{m_i}$ [€/h] the average malfunction costs of the overburden excavation system caused by failures of belt conveyor rubber belt of length L_i.

T_{pm} = 12 [h] the time needed for the planned replacement of the rubber belt (gradual failures).

T_{npm} = 30 [h] the time needed for the unplanned replacement of the rubber belt (sudden failures).

$c_{pm} = T_{pm} \cdot c_{m/h}$ = 12·9232.33 = 110787.96 [€] – the malfunction costs of the overburden excavation system with the planned replacement of the rubber belt.

$c_{npm} = T_{npm} \cdot c_{m/h}$ = 30·9232.33 = 276969.90 [€] – the malfunction costs of the overburden excavation system with the unplanned replacement of the rubber belt.

$c_{m/h}$ = 9232.33 [€/h] – the malfunction costs of the overburden excavation system per hour – see expression (6).

$MTTF_{gf}^{L_i}$ [h] the mean time to gradual failure of the rubber belt of length L_i, (Tables 2, 3, 4, 5, 6, 7).

$H(MTTF_{gf}^{L_i})$ the renewal function of the rubber belt of length (L_i), in the period of $MTTF_{gf}^{L_i}$ i.e. for the mean time to gradual failures.

$c_{nbr}^{L_i}$ [€] the unused rubber belt resource of length L_i.
In another words, expression (5) gives us the unit cost [€/h] of the overburden excavation system malfunction caused by belt conveyor rubber belt failure in the period of $MTTF_{gf}^{L_i}$ [h]. In that period one planned rubber belt replacement and $H(MTTF_{gf}^{L_i})$ unplanned rubber belt replacements will be carried out.

The malfunction costs of the overburden excavation system per hour $c_{m/h}$ [€/h], expressed through the price of the final product – lignite, can be calculated as:

$$c_{m/h} = Q_T \cdot \bar{k}_t \cdot \bar{k}_q \cdot r_e \cdot c_{l/t} = 4100 \cdot 0.466 \cdot 0.464 \cdot 0.833131 \cdot 12.5 =$$

$$= 9232.33 \text{ [€/h]} \qquad (6)$$

where:

Q_T =4100 [m³lm/h] – the theoretical digging capacity of the bucket wheel excavator, cubic meters of loose material per hour.

$\bar{k}_t$ = 0.466 – the average coefficient of time utilisation (Table 1.).

$\bar{k}_q$ = 0.464 – the average coefficient of capacity utilisation (Table 1.).

$r_e = \dfrac{Q_l^{tot}}{Q_{ob}^{tot}} = \dfrac{140619324}{168784098.6} = 0.833131$ [t/m³lm] – the ratio of overburden quantity which has to be removed (cubic meters of loose material) in order to excavate one ton of lignite (Table 1).

$c_{l/t}$ = 12.5 [€/t] – lignite price per ton (in Serbia in the last 20 years the price of electrical energy is a social category, therefore this lignite price is not real, i.e. the market price, because it is determined as a percentage of the price of one kWh of electrical energy).[2]

Rubber belt working time until sudden failure can be described by exponential distribution (see previous chapter), which means that the renewal function of the rubber belts has the following form:

$$H(MTTF_{gf}^{L_i}) = \lambda_i \cdot MTTF_{gf}^{L_i} \qquad (7)$$

where:

$\lambda_i = 1/MTTF_{sf}^{L_i}$ [1/h] – the sudden failure intensity of the rubber belt of length L_i, and

$MTTF_{sf}^{L_i}$ [h] – the mean (working) time to sudden failure of the rubber belt of length L_i, (Table 10).

In the case of gradual failures, it is assumed that the complete rubber belt resource is used, while in the case of sudden failures one part of the rubber belt resource remains unused. The value of the unused rubber belt resource of length L_i, can be determined from the following expression:

$$c_{nbr}^{L_i} = c_{b/m} \cdot L_i \cdot \left(1 - \frac{MTTF_{sf}^{L_i}}{MTTF_{gf}^{L_i}}\right), \text{ [€]} \qquad (8)$$

where:

$c_{b/m}$ = 200 [€/m] – the rubber belt price (B1600) per meter [13].

The average malfunction costs (lost production) of the overburden excavation system caused by belt conveyor rubber belt failure, i.e. the unit cost of system malfunction per hour of belt conveyor work during belt lifetime, as a function of rubber belt length L_i are shown in Table 11 and Figure 4.

Figure 4 shows that the average malfunction costs of the overburden excavation system $\bar{C}_{m_i}$ per hour of belt conveyor work during belt lifetime decrease with the increase of rubber belt length L_i. Figure 4 shows that costs $\bar{C}_{m_i}$, depending on rubber belt length, range between 100 and 200 €/h.

Table 11.Average malfunction costs of the overburden excavation system $\bar{C}_{m_i}$ [€/h].

i	L_i [m]	λ_i [1/h]	$H(MTTF_{gf}^{L_i})$	$c_{npm} + c_{nbr}^{L_i}$ [€]	$(c_{npm} + c_{nbr}^{L_i}) \cdot H(MTTF_{gf}^{L_i})$ [€]	$\bar{C}_{m_i}$ [€/h]
1	27.6	0.000521	1.130187	277605.59	313746.36	195.84
2	51.6	0.000313	0.898776	276969.74	248933.86	125.38
3	75.0	0.000523	2.684894	286382.92	768907.85	171.27
4	81.5	0.000381	1.957722	284943.73	557840.63	130.21
5	106.1	0.000258	1.394951	282977.73	394739.98	93.32
6	158.2	0.000421	4.135450	300958.81	1244600.12	137.88

[2] Source: Public company Electric Power Industry of Serbia.

4. Conclusion

For large mining systems like overburden excavation systems, it is important to determine when the equipment in a system will breakdown, or how long it will perform in a reliable manner in order to take the necessary precautions to ensure continuity of the system operation due to huge lost production costs. Therefore the proposed methodology and its output –average malfunction (lost production) costs are designed to assist persons in charge in planning and adopting adequate maintenance strategies based on the prediction of failure intensities and lost production costs.

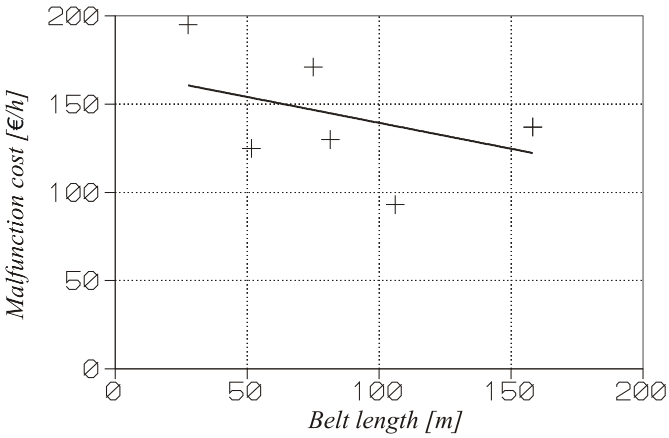


Fig. 4. Average malfunction costs

Statistical processing and analysis of the data pertaining to the malfunctions of the overburden excavation system caused by rubber belt failures taken from the maintenance diary (plant records) of the Tamnava – East field open-pit mine using the proposed methodology lead to the following conclusions.

The working time of the belt conveyor rubber belts, which work in the overburden excavation system (bucket wheel excavator, belt wagon, spreader) can be described by different theoretical distributions of time due to the nature of failure. In another words, rubber belt working time to failure presents a complex mathematical model, in which working time until sudden failure (tear, breakthrough) can be described by exponential distribution while working time until gradual failure can be described by normal distribution.

It is also shown that there is a strong linear dependence between mean time to failure, caused by gradual failures and rubber belt length. The obtained dependence indicates that the proposed methodology of analysis of rubber belt working time to failure can be applied to other rubber belts of different lengths working on similar machines.

The analysis of rubber belt working time to failure using the proposed methodology as a complex mathematical model enables the calculation of the average malfunction costs of the overburden excavation system (lost production) caused by belt conveyor rubber belt failure, i.e. the unit cost of system malfunction per hour of belt conveyor work during belt lifetime. Malfunction costs determined in this way facilitate better planning of malfunctions, requirements for spare rubber belts as well as a reduction of working costs in the open-pit mine.

The obtained results using the proposed methodology of analysis of belt conveyor rubber belt working time to failure are valid only for working conditions on the Tamnava – East field open-pit mine, while the proposed methodology as well as the expression for malfunction cost determination can be applied, with appropriate adoptions, in the analysis of other open-pit mines in order to indicate better (optimal) maintenance strategy.

References

1. Baldin A, Furlanetto L, Roversi A, Turco F. Manual for maintenance of industrial facilities. Belgrade: Maintenance of machines and equipments – OMO, 1979.
2. Barabady J, Kumar U. Reliability analysis of mining equipment: A case study of a crushing plant at Jajarm Bauxite Mine in Iran. Reliability engineering & system safety 2008; 93(4): 647-653.
3. Bindzar P, Grincova A, Ristovic I. 3D Mathematical Model of Conveyor Belt Subjected to a Stress Loading. Underground Mining Engineering 2006; 13(15): 81-89.
4. Chookah M, Nuhi M, Modarres M, Seibi A. Structuring a probabilistic model for reliability evaluation of piping subject to corrosion-fatigue degradation. American Nuclear Society - International Topical Meeting on Probabilistic Safety Assessment and Analysis PSA 2008. 2008; 3: 1718-1740.
5. Elevli S, Demirci A. The effects of maintenance activities on reliability and profitability. CIM Bulletin 1999; 92(1034): 49-51.
6. Hardygora M. Trends in Conveyor belt Research. Transport & Logistics 2002; 3: 1-12.
7. Hoseinie S H, Ataei M, Khalokakaie R, Kumar U. Reliability modeling of water system of longwall shearer machine. Archives of mining sciences 2011; 56(2): 291-302.
8. Huang L P, Yue W H, Gong Z L. Reliability modeling and simulation of mechanical equipment undergoing maintenance. Applied Mechanics and Materials 2010; 34-35: 1211-1216.
9. Kumar U. Special issue on reliability and maintenance of mining systems. International Journal of Mining Reclamation and Environment 2009; 23(3): 155-156.
10. Liu D, Huang L, Yue W, Xu X. Reliability simulation and design optimization for mechanical maintenance. Chinese Journal of Mechanical Engineering 2009; 22(4): 594-601.
11. Nenadovic M. Mathematical processing of measured data. Belgrade: Serbian academy of science and arts, 1988.
12. Petrovic D, Bugaric U. Analysis of stochastic values as a part of transport system design. Proc. of 1st Int. Symposium of Industrial Engineering. Belgrade 1996; 411-413.
13. Polovina D, Ignjatovic D, Ristovic I. Optimal time for replacement of rubber belts in excavator conveyors for excavation waste. Proc. of 22nd Yugoslav symposium of operations research SYM-OP-IS ’95. Donji Milanovac 1995; 551-554.
14. Tabachnik B G, Fidell L S. Using Multivariate Statistics. Boston: Pearson Education Inc., 2007.
15. Uzgoren N, Elevli S, Elevli B, Uysal O. Reliability analysis of draglines' mechanical failures. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2010; 48: 23-28.
16. Uzgoren N, Elevli S. Nonhomogeneous Poisson process: reliability analysis of a mining equipment. Journal of the faculty of engineering and architecture of Gazi university 2010; 25(4): 827-837.
17. Yue W, Wang X, Huang L. Cost modeling and simulation for mining mechanical equipment undergoing maintenance. Advanced Materials Research 2012; 418-420: 2122-2125.

Associate professor Ugljesa BUGARIC
University of Belgrade
Faculty of Mechanical Engineering
Kraljice Marije 16, 11120 Belgrade 35, Serbia
E-mail: ubugaric@mas.bg.ac.rs

Assistant professor Milos TANASIJEVIC
University of Belgrade
Faculty of Mining and Geology
Djusina 7, 11000 Belgrade, Serbia
E-mail: tan@rgf.bg.ac.rs

Dragan POLOVINA, Ph.D., Eng.
Public company Electric power industry of Serbia
Vojvode Stepe 412, 11000 Belgrade, Serbia
E-mail: dragan.polovina@eps.rs

Professor Dragan IGNJATOVIC
University of Belgrade
Faculty of Mining and Geology
Djusina 7, 11000 Belgrade, Serbia
E-mail: gagi@rgf.bg.ac.rs

Assistant professor Predrag JOVANCIC
University of Belgrade
Faculty of Mining and Geology
Djusina 7, 11000 Belgrade, Serbia
E-mail: pjovancic@rgf.bg.ac.rs

Article citation info:
SU C, ZHANG YJ, CAO BX. Forecast model for real time reliability of storage system based on periodic inspection and maintenance data. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 342–348.

Chun SU
Yong-jin ZHANG
Bai-xue CAO

FORECAST MODEL FOR REAL TIME RELIABILITY OF STORAGE SYSTEM BASED ON PERIODIC INSPECTION AND MAINTENANCE DATA

MODEL DO PROGNOZOWANIA NIEZAWODNOŚCI SYSTEMU MAGAZYNOWANIA W CZASIE RZECZYWISTYM W OPARCIU O DANE Z PRZEGLĄDÓW OKRESOWYCH ORAZ DANE EKSPLOATACYJNE

In recent years, storage reliability has attracted much attention for increasing reliability requirement. In this paper, forecast models for real-time reliability of storage system under periodic inspection and maintenance are presented, which is based on the theories of reliability physics and exponential distribution. The models are developed under two newly-defined imperfect repair modes, i.e., Improved As Bad As Old (I-ABAO) and Improved As Good As New (I-AGAN). A completion method for censored life data is also proposed by averaging the residual lifetime. According to the complete and censored lifetime data, parameters in the models are estimated by applying maximum likelihood estimation method and iterative method respectively. A numerical example of a storage system is given to verify the feasibility of the proposed completion method and the effectiveness of the two models.

Keywords: storage reliability, real-time reliability, periodic inspection and maintenance, censored data.

W ostatnich latach wiele uwagi poświęcono tematyce niezawodności magazynowania w odniesieniu do zwiększania wymogu niezawodności. W prezentowanym artykule, przedstawiono modele do prognozowania w czasie rzeczywistym niezawodności systemu magazynowania podlegającego przeglądom okresowym i obsłudze. Modele oparto na teoriach z zakresu fizyki niezawodności oraz na rozkładzie wykładniczym. Proponowane modele opracowano dla dwóch nowo zdefiniowanych opcji niepełnej odnowy, t.j. Improved-As Bad As Old (Jak Tuż Przed Uszkodzeniem – Wersja Udoskonalona) oraz Improved-As Good As New (Jak Fabrycznie Nowy – Wersja Udoskonalona). Zaproponowano także metodę uzupełniania danych cenzurowanych (uciętych) dotyczących trwałości polegającą na uśrednianiu trwałości resztkowej. Zgodnie z pełnymi i cenzurowanymi danymi trwałościowymi, parametry w proponowanych modelach ocenia się, odpowiednio, z zastosowaniem estymacji metodą największej wiarygodności oraz metody iteracyjnej. Poprawność przedstawionej metody uzupełniania oraz efektywność proponowanych dwóch modeli zweryfikowano na podstawie numerycznego przykładu systemu magazynowania.

***Słowa kluczowe**: niezawodność magazynowania, niezawodność w czasie rzeczywistym, okresowe przeglądy i obsługa, dane cenzurowane.*

1. Introduction

In engineering practice, there exit such kinds of systems. They are kept in storage state for most of their lifetime, but must keep high mission reliability and be ready for operation whenever needed. Strategic missile, spare parts for nuclear power plant, and fire protection system are typical examples of such systems. Periodic testing, replacement and other maintenance measures are necessary to avoid or reduce the occurrence of failure [6, 9, 12].

In the past decades, storage reliability has attracted much attention from both academia and engineering field. Merren studied the periodic test problem of electronic equipment in storage, and an algorithm for computing the reliability was developed based on test data [8]. Aneziris presented a method for calculating the dynamic reliability of safety systems, which was based on the theory of Markov chain [2]. Ref. [4] proposed an availability model for storage products under periodic inspection, but they ignored the initial failure. Zhao and Xie studied the problem of potential storage reliability estimation with initial storage failure [13]. Zhang and Zhao established a storage reliability model under periodic inspection [14]. Ref. [15] studied the storage reliability of missile-engine with Bayesian approach. Yang used stochastic process method to describe the degradation process of the components with multi-performance parameters, and proposed a storage reliability evaluation model [11]. Ref. [13] introduced an approach to evaluate the storage reliability of engine control circuit module.

Up to now, most existing research mainly concentrated on the data analysis methods for the components of a system. Real-time reliability estimation of products in storage on the system level has not received sufficient attention. Most studies assume the system is in perfect initial state and there is not much effort placed on empirical study based on actual storage reliability data. As for repair efficiency, two basic assumptions are known As Bad As Old (ABAO) and As Good As

New (AGAN). However, these two extreme cases seldom happen in most practical systems [7]. More reasonable repair models need to be developed to get a closer description of the real situation.

In this paper, real-time reliability evaluation models are established based on periodic inspection data, and two newly-defined imperfect repair modes are considered, i.e., Improved-ABAO and Improved-AGAN. A new completion method is also proposed to convert censored life data into complete data. Based on complete and censored lifetime data, the parameters in the models are estimated with maximum likelihood estimation method and iterative method, respectively. A numerical example is given to illustrate the models and method.

2. Reliability model for storage system

Due to different reasons, not all the storage systems are shipped out of the factory with reliability testing. Thus the initial use reliability of the system might not be 1.

Let A indicate the event "the storage system is qualified at t=0", denote T as the lifetime of the storage system. The real-time reliability at time $t(t\geq 0)$ can be expressed as:

$$R(t)=p(T>t,A)=p(T>t\mid A)p(A)=R_0R_s(t) \tag{1}$$

where R_0 is the inherent reliability of a storage system, $R_0=p(A)$; $R_s(t)=p(T>t\mid A)$ is the conditional reliability of the system, which changes with storage time and $R_s(0)=1$. Therefore, the system reliability can be represented as

$$R(t;\theta)=R_0R_S(t;\theta),\theta\in\Theta \tag{2}$$

where Θ is a measurable parameter set.

In order to keep high mission reliability, periodic inspection and maintenance are needed. Based on the difference of maintenance, two types of forecast models for real-time reliability are given as follows.

2.1. Reliability forecast model under I-AGAN

Assume that the storage system is put into use at t=0 and it is inspected at intervals of time $\tau(\tau\geq 0)$. If partial failure or aging occurs, the maintenance is carried out. After $x(0<x<\tau)$ time maintenance, the system is restored and keeps on storing. When the maintenance is perfect and leaves the system as if it were new, we call it repair under AGAN[7]. In engineering practice, however, maintenance can increase system reliability but the failure rate usually shows an increasing tendency owing to the degradation of material strength. Taking account of this, we propose a new repair mode called Improved-AGAN (I-AGAN) (See Fig. 1).

Under I-AGAN, three assumptions are made as below: (1) The system is stored under natural conditions, and the lifetime t obeys exponential distribution. (2) In the initial time, the system reliability is $R(0)=R_0\leq 1$ and lifetime t obeys exponential distribution with parameter θ_0. After the k[th] maintenance, lifetime obeys exponential

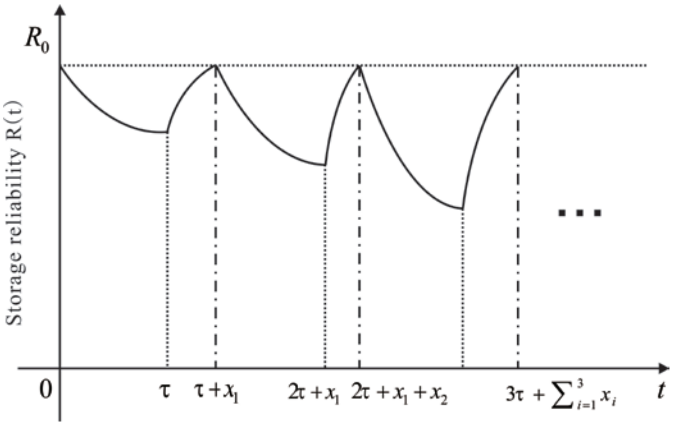


Fig. 1 Reliability change of storage system under I-AGAN

distribution with parameter θ_k, and $\theta_0\geq\theta_1\geq\theta_2\geq\ldots\geq\theta_{k-1}\geq\theta_k$. (3) The inspection interval is τ. The maintenance time x_k (k = 1, 2, …) is a stochastic variable that varies with failure.

From Fig. 1, the reliability before the k[th] maintenance $k\tau+\sum_{i=1}^{k-1}x_i$ is $R_s(k\tau+\sum_{i=1}^{k-1}x_i)$. According to the definition under I-AGAN, after each maintenance the system is restored to its initial state, so the reliability at time $k\tau+\sum_{i=1}^{k-1}x_i$ is $R_s(k\tau+\sum_{i=1}^{k-1}x_i)\leq R_s(k\tau+\sum_{i=1}^{k}x_i)=R_0=R(0)$, where k=1, 2, 3, …, n; n is times of maintenance.

Denote $F(t)$ as the failure distribution function of the system at t, and $F(t)$=1-R(t). Hence, $F(t)$ obeys the following distribution:

$$F(t)=1-R_0\cdot R_s\left(t-r\tau-\sum_{i=1}^{r}x_i\right),r=\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right] \tag{3}$$

where [·] represents rounding down.

Based on the characteristic of exponential distribution, in non-maintenance period $R(t)$ can be calculated:

$$R(t)=R_0\exp\left(-\frac{t-r\tau-\sum_{i=1}^{r}x_i}{\theta_r}\right),r=\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right] \tag{4}$$

Suppose λ_i (i=1, 2, …, k, …) is the system failure rate after the i[th] maintenance. As $\lambda_i=1/\theta_i$, obviously, $\lambda_0\leq\lambda_1\leq\lambda_2\leq\ldots\leq\lambda_{k-1}\leq\lambda_k$. According to the reliability physics and engineering practice, after the k[th] maintenance the failure rate satisfies the following equation:

$$\lambda_k=\lambda_0\exp(k\beta),\ k=1,2,\cdots,n\,, \tag{5}$$

where $\beta(\beta>0)$ is the degradation factor. It can also be rewritten as $\theta_k=\theta_0\exp(-k\beta)$, k = 1, 2, …, n.

From Eqs. (4), (5), the real-time reliability at time t is:

$$R(t)=R_0\exp\left[-\lambda_0\exp(r\beta)(t-r\tau-\sum_{i=1}^{r}x_i)\right],r=\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right] \tag{6}$$

Based on the maintenance strategy, it can be obtained that:

$$\lim_{t\to(k\tau+\sum_{i=1}^{k-1}x_i)^-}\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right]=k-1\,,$$

$$\lim_{t\to(k\tau+\sum_{i=1}^{k}x_i)}\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right]=k,k=1,2,\cdots,n \tag{7}$$

From Eq. (6), it can be concluded that the system real-time reliability before the kth maintenance time $t=k\tau+\sum_{i=1}^{k-1}x_i$ is:

$$R_k^-(t)=R_0\exp(-\lambda_0e^{k\beta}\tau),k=1,2,\cdots,n \tag{8}$$

2.2. Reliability forecast model under I-ABAO

When the maintenance leaves the system in the same state as it was before failure, we call it repair under ABAO[7]. In practice, however, maintenance may leave the system reliability higher than it was before failure but lower than its initial state R_0. Additionally, material strength degradation causes failure rate to increase. Taking account of this, we

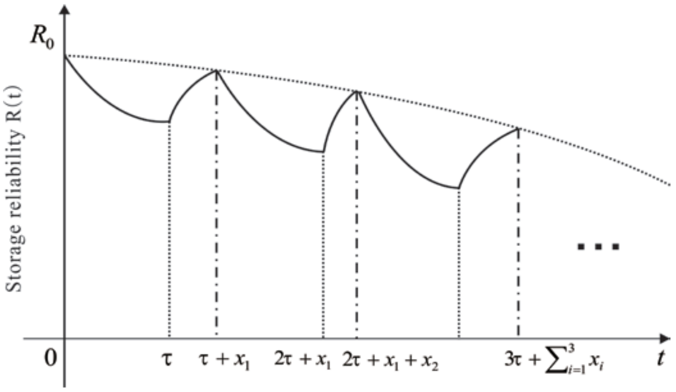


Fig. 2 Reliability change of storage system under I-ABAO

propose another imperfect repair mode called Improved-ABAO (I-ABAO), with the same assumptions in Section 2.1 (See Fig. 2).

According to reliability characteristics, we can divide a storage system into two subsystems. Subsystem 1 requires regular inspection and maintenance with an invariable or increasing failure rate. Subsystem 2 has a high reliability and does not require regular inspection and maintenance, but aging phenomenon may occur.

Under I-ABAO, after each maintenance the system reliability $R_s(t)$ cannot be restored to the initial reliability, that is: $R_s(k\tau+\sum_{i=1}^{k}x_i)<R_0=R(0)$. During the storage, the real-time reliability of Subsystem 1 can be obtained by:

$$R_{s1}(t)=\exp\left[-\lambda_0\exp(r\beta)(t-r\tau-\sum_{i=1}^{r}x_i)\right],\ r=\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right] \quad (9)$$

Subsystem 2 is composed of mechanical and electronic parts, which has usually undergone ageing screening tests, and the infancy failures have been eliminated. Basically, it is in the random failure period, so the system reliability is assumed to decrease exponentially, which is expressed as[1]:

$$R_{s2}(t)=\exp(-\delta t) \quad (10)$$

where δ is the degradation coefficient. From Eq. (2), it can be concluded that the system reliability at time t is:

$$R(t)=R_0R_s(t)=R_0R_{s1}(t)R_{s2}(t)=R_0\exp\left[\lambda_0\exp(r\beta)\left(r\tau+\sum_{i=1}^{r}x_i-t\right)-\delta t\right],$$

$$r=\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right] \quad (11)$$

Before the k^th^ inspection, which is at time $t=k\tau+\sum_{i=1}^{k-1}x_i$, the system reliability $R_k^-(t)$ is:

$$R_k^-(t)=R_0\exp\left\{-\left[\lambda_0\exp(r\beta)+\delta(k\tau+\sum_{i=1}^{k-1}x_i)\right]\right\},\ k=1,2,\cdots,n \quad (12)$$

3. Completion of censored data

Regular inspection is necessary for the storage system, but the failures are seldom found exactly at the inspection time. So it is difficult to obtain complete data, and the field lifetime data are mostly censored. If the censored data are simply regarded as complete data or treated with interpolation method, large errors will be brought out. The completion method of censored life data is discussed below.

3.1. Completion method of censored data

Let $T_1, T_2, \ldots, T_n$ be the lifetimes of storage system; $t_1, t_2, \ldots, t_n$ be the start times of inspection and $o(T_i)$ be the residual lifetime of the i^th^ system. If t_i is right censored or completely censored, $I_i = 0$; Otherwise, if t_i is left censored, $I_i=1$. It can be obtained that $o(T_i)=(T_i-t_i)(1-I_i)$, $I_i=I\{t_i\le T_i\}, i=1,2,\cdots,n$. Consequently, the complete data of $t_i(i=1,2,\cdots,n)$ is:

$$\xi_i=t_i+o(T_i) \quad (13)$$

Accordingly, when the failure data of the i^th^ system is right censored, the average residual lifetime is:

$$E(T_i-t_i\mid T_i\ge t_i)=\frac{p\{T_i-t_i,T_i\ge t_i\}}{p\{T_i\ge t_i\}}=\frac{\int_{t_i}^{+\infty}(x-t_i)f(x;\theta_1,\theta_2,\cdots,\theta_m)dx}{1-F_s(t_i;\theta_1,\theta_2,\cdots,\theta_m)} \quad (14)$$

Therefore, under periodic inspection the average residual lifetime of the system is:

$$\overline{o(T_i)}=(1-I_i)\frac{\int_{t_i}^{+\infty}(x-t_i)f(x;\theta_1,\theta_2,\cdots,\theta_m)dx}{R_s(t_i;\theta_1,\theta_2,\cdots,\theta_m)} \quad (15)$$

If inspection data are discontinuous, the average residual lifetime is:

$$D\overline{o(T_i)}=\frac{(t_{i+1}-t_i)R_s(t_i)+\sum_{j\ge i+1}(t_{j+1}-t_j)R_s(t_j)}{R_s(t)} \quad (16)$$

where the measurable parameter set $\theta_1,\theta_2,\cdots,\theta_m\in\Theta$ and $\overline{o(T_i)}$ are the implicit functions of unknown parameters, which can be solved by iterative algorithm [10]. Assuming that the system lifetime T has density function $f(\xi;\Theta)$, the likelihood function of parameter set Θ within the complete sample data $\xi_1,\xi_2,\cdots,\xi_n$ is $L(\xi;\Theta)=\prod_{i=1}^{n}f(x_i;\Theta)$, so the likelihood equation is:

$$\left.\frac{\partial\ln L(\xi;\theta_1,\theta_2,\cdots,\theta_m)}{\partial\theta_i}\right|_{\theta_i-\hat{\theta}_i}=0, i=1,2,\cdots,m \quad (17)$$

Suppose the initial value of parameter estimation is $\Theta_0=\{(\theta_1)_0,(\theta_2)_0,\cdots,(\theta_m)_0\}$. Insert it into Eq. (17) and the first iterative parameter set $\Theta_1=\{(\theta_1)_1,(\theta_2)_1,\cdots,(\theta_m)_1\}$ can be obtained. Circulating the iterative procedure, the kth iterative parameter set $\Theta_k=\{(\theta_1)_k,(\theta_2)_k,\cdots,(\theta_m)_k\}$ can also be obtained. Thereby, the maximum likelihood estimator set of Θ is the convergent parameter set $\hat{\Theta}=\lim_{k\to\infty}\Theta_k$.

3.2. Censored data completion of Weibull & exponential type system

For most storage system, the lifetime obeys Weibull distribution or memoryless exponential distribution. Based on cumulative damage effect, the method to convert censored data of Weibull storage system into complete ones is given below. Assume that the lifetime T has the density function:

$$f(t;\alpha,\lambda)=\lambda\alpha(\lambda t)^{\alpha-1}\exp\left[-(\lambda t)^{\alpha}\right],\ t\ge 0 \quad (18)$$

where λ, α are parameters to be estimated. The likelihood function under the case of the complete sample is:

$$L(T_1,T_2,\cdots,T_n;\lambda,\alpha)=\prod_{i=1}^{n}\lambda\alpha(\lambda T)^{\alpha-1}\exp\left[-\left(\lambda T_i\right)^{\alpha}\right]. \quad (19)$$

Therefore, the logarithmic likelihood equations are:

$$\begin{cases}\dfrac{\sum_{i=1}^{n}T_i^{\alpha}\ln T_i}{\sum_{i=1}^{n}T_i^{\alpha}}-\dfrac{1}{\alpha}-\dfrac{1}{n}\sum_{i=1}^{n}\ln T_i=0\\ \lambda^{-\alpha}-\dfrac{1}{n}\sum_{i=1}^{n}T_i^{\alpha}=0\end{cases} \quad (20)$$

If x_i is left censored, it can be obtained from Eqs. (13), (15) that:

$$T_i=\frac{\lambda\int_{x_{i-1}}^{x_i}t^{\frac{1}{\alpha}}\exp(-t)dt}{\exp(-\pi_{i-1})-\exp(-\pi_i)},\ \pi_{i-1}=\left[\lambda\left((i-1)\tau+\sum_{j=1}^{i-1}x_j\right)\right]^{\alpha},\ \pi_i=\left[\lambda\left(i\tau+\sum_{j=1}^{i-1}x_j\right)\right]^{\alpha} \quad (21)$$

If x_i is right censored, then:

$$T_i=\frac{1}{\lambda}\exp\left[\left(\lambda x_i\right)^{\alpha}\right]\left[\Gamma\left(1+\frac{1}{\alpha}\right)-\int_0^{(\lambda x_i)^{\alpha}}t^{\frac{1}{\alpha}}\exp(-t)dt\right] \quad (22)$$

If x_i is completely censored, take T_i as:

$$T_i=x_i \quad (23)$$

For the exponential storage system, take $\alpha = 1$, then the completion of left censored data and right censored data are given as follows, respectively.

$$T_i=\frac{1}{\lambda}+\frac{\tau}{\exp(\lambda\tau)-1}+\left[(i-1)\tau+\sum_{j=1}^{i-1}x_j\right] \text{ and } T_i=x_i+\frac{1}{\lambda} \quad (24)$$

4. Parameter estimation of reliability model

When the censored data have been converted into complete data, the estimator $\hat{\Theta}$ of parameter set can be obtained by iteration of Eq. (17). Given the distribution function $F(t)=F(t;\hat{\Theta})$, the real-time reliability function of storage system can be obtained as $R(t)=\hat{R}_0\cdot[1-F(t;\hat{\Theta})]$.

The least square method was used to estimate the initial reliability in Ref. [13]. Here we take

$$\hat{R}_0=\min(R_0',\frac{R_0'+\hat{p}_0}{2}) \quad (25)$$

as the initial reliability of the system, where $\hat{p}_0=(N_0-f_0+1)/(N_0+2)$ is the Bayes estimation of use reliability, N_0 is the number of systems, and f_0 is the failure number. For Weibull storage system, the real-time reliability function is:

$$R(t)=\hat{R}_0\cdot\exp\left[-(\hat{\lambda}t)^{\hat{\alpha}}\right] \quad (26)$$

For the original censored data, $\{(N_i, S_i, t_i), i=1, 2, \ldots, n\}$ is the failure data set collected from periodic inspection, where S_i is the number of non-failed products within N_i of total stored products at time t_i. Supposing that the periodic inspection interval τ is a fixed time unit, and the maintenance time is negligible compared to the storage time, only parameters in Eqs. (8), (12) need to be estimated.

For real-time reliability model under I-AGAN:

$$R_k^-(t)=R_0\exp\left[-\lambda_0\exp(k\beta)\right],\ k=1,2,\cdots,n \quad (27)$$

The likelihood function is:

$$L_1(\lambda,\beta\,|\,\Theta)=\prod_{i=1}^{n}\binom{N_i}{S_i}\cdot\left[R_0\exp(-\lambda_0 e^{i\beta})\right]^{S_i}\times\left[1-R_0\exp(-\lambda_0 e^{i\beta})\right]^{N_i-S_i} \quad (28)$$

The maximum likelihood equations involving λ_0 and β are as follows:

$$\begin{cases}\dfrac{\sum_{i=1}^{n}S_i e^{i\beta}}{R_0}-\sum_{i=1}^{n}\dfrac{(N_i-S_i)e^{i\beta}\exp(-\lambda_0 e^{i\beta})}{1-R_0\exp(-\lambda_0 e^{i\beta})}=0\\ \dfrac{\sum_{i=1}^{n}iS_i e^{i\beta}}{R_0}-\sum_{i=1}^{n}\dfrac{ie^{i\beta}(N_i-S_i)\exp(-\lambda_0 e^{i\beta})}{1-R_0\exp(-\lambda_0 e^{i\beta})}=0\end{cases} \quad (29)$$

It can be concluded from Eq. (29) that there is no analytical solution for maximum likelihood estimation of λ_0 and β. However, the estimates $\hat{\lambda}$, $\hat{\beta}$ can be derived by numerical calculating method. The real-time system reliability is:

$$R(t)=\hat{R}_0\exp\left[-\hat{\lambda}_0\exp(r\hat{\beta})(t-r\tau-\sum_{i=1}^{r}x_i)\right],\ r=\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right] \quad (30)$$

For the real-time reliability model under I-ABAO:

$$R_k^-(t)=R_0\exp\left\{-\left[\lambda_0\exp(r\beta)+\delta(k\tau+\sum_{i=1}^{k-1}x_i)\right]\right\} \quad (31)$$

The corresponding maximum likelihood function is:

$$\begin{aligned}L_2(\lambda_0,\delta,\beta\,|\,\Theta)=&\prod_{k=1}^{n}\binom{N_k}{S_k}\cdot\left\{R_0\exp\left[-\left(\lambda_0 e^{k\beta}+\delta(k+\sum_{i=1}^{k-1}x_i)\right)\right]\right\}^{S_k}\times\\&\left\{1-R_0\exp\left[-\left(\lambda_0 e^{k\beta}+\delta(k+\sum_{i=1}^{k-1}x_i)\right)\right]\right\}^{N_k-S_k}\end{aligned} \quad (32)$$

Then the maximum likelihood equations involving λ, δ, β are as follows:

$$\begin{cases}\dfrac{\sum_{k=1}^{n}S_k e^{k\beta}}{R_0}-\sum_{k=1}^{n}\dfrac{(N_k-S_k)e^{k\beta}\exp\left[-\left(\lambda_0 e^{k\beta}+(k+\sum_{i=1}^{k-1}x_i)\delta\right)\right]}{1-R_0\exp\left[-\left(\lambda_0 e^{k\beta}+(k+\sum_{i=1}^{k-1}x_i)\delta\right)\right]}=0\\ \dfrac{\sum_{k=1}^{n}kS_k e^{k\beta}}{R_0}-\sum_{k=1}^{n}\dfrac{ke^{k\beta}(N_k-S_k)\exp\left[-\left(\lambda_0 e^{k\beta}+(k+\sum_{i=1}^{k-1}x_i)\delta\right)\right]}{1-R_0\exp\left[-\left(\lambda_0 e^{k\beta}+(k+\sum_{i=1}^{k-1}x_i)\delta\right)\right]}=0\\ \dfrac{\sum_{k=1}^{n}kS_k}{R_0}-\sum_{k=1}^{n}\dfrac{(N_k-S_k)k\exp\left[-\left(\lambda_0 e^{k\beta}+(k+\sum_{i=1}^{k-1}x_i)\delta\right)\right]}{1-R_0\exp\left[-\left(\lambda_0 e^{k\beta}+(k+\sum_{i=1}^{k-1}x_i)\delta\right)\right]}=0\end{cases} \quad (33)$$

From Eq. (33), the estimates $\hat{\lambda}_0$, $\hat{\delta}$, $\hat{\beta}$ of λ_0, δ, β can be derived by Newton's iteration method. The real-time system reliability is:

$$R(t)=\hat{R}_0\exp\left[\hat{\lambda}_0\exp(r\hat{\beta})(r\tau+\sum_{i=1}^{r}x_i-t)-\hat{\delta}\,t\right],\ r=\left[\left(t-\sum_{i=1}^{r}x_i\right)/\tau\right] \quad (34)$$

5. Case study

In this section, a storage system composed of 18 subsystems is considered. It is maintained every other year to keep its availability. Compared with the storage time, the time consumed by maintenance is negligible. Those subsystems which have serious failure or have no value for maintenance will be withdrawn from storage, while others will be restored to storage state. Maintenance will continue until all the subsystems are withdrawn from storage. The detailed data are shown in Table 1.

Table 1. Recorded data of 18 subsystems under periodic inspection

Year (t_i)	Total storage (N_i)	Non-failures (S_i)	Cumulated withdrawals (F_i)
1	18	18	0
2	18	18	0
3	18	16	2
4	16	16	2
5	16	16	2
6	16	16	2
7	16	13	5
8	13	12	6
9	12	12	6
10	12	11	7
11	11	10	8
12	10	9	9
13	9	4	14
14	4	4	14
15	4	3	15
16	3	3	15
17	3	1	17
18	1	1	17
19	1	1	17
20	1	0	18

The reliability analysis of the system under I-AGAN and I-ABAO is showed as below.

5.1. Reliability function of Weibull & exponential type system

At the beginning of storage, all the 18 subsystems satisfy mission reliability. According to Eq. (25), the initial storage reliability R_0 is:

$$\hat{R}_0 = \min\left(1, \frac{1}{2}(1+\frac{S_0+1}{N_0+2})\right) = 0.975$$

Rearrange the data in Table 1 on a monthly basis by left-censored data (denoted as x⁻), right-censored data (denoted as x⁺), and complete data (denoted as x): { 36+, 36+, 84+, 84+, 84+, 96+, 120+, 132-, 144+, 156+, 156+, 156+, 156+, 156, 180+, 204+, 204+, 228+ }.

If the storage lifetime obeys Weibull distribution, the complete lifetime data of the storage system T_i (month) are derived from Eqs. (21), (22), (23), as shown in the sequence: {248, 248, 262, 262, 262, 267, 278, 126, 291, 298, 298, 298, 298, 156, 312, 327, 327, 343}. Inserting it into Eq. (20), parameter estimators with a precision of 10-5 are obtained: $\hat{\alpha}$=2.08998, $\hat{\lambda}$=0.00363. The real-time reliability function with complete lifetime data is:

$$R(t) = 0.975\exp\left[-(0.00363t)^{2.08998}\right] \quad (35)$$

If the storage lifetime obeys exponential distribution, the complete lifetime data T_i (month) are derived from Eq. (24), as shown in the sequence: {1239, 1239, 1287, 1287, 1287, 1299, 1323, 126, 1347, 1359, 1359, 1359, 1359, 156, 1383, 1407, 1407, 1431}. The estimator of parameter λ is:

$$\hat{\lambda} = \left(\frac{1}{n}\sum_{i=1}^{n} T_i\right)^{-1} = 8.3126\times 10^{-4}.$$

So the real-time reliability function with the complete lifetime data is:

$$R(t) = 0.975\cdot\exp\left(-0.00083126t\right) \quad (36)$$

Fig. 3 illustrates the reliability change of Weibull type system and exponential type system.

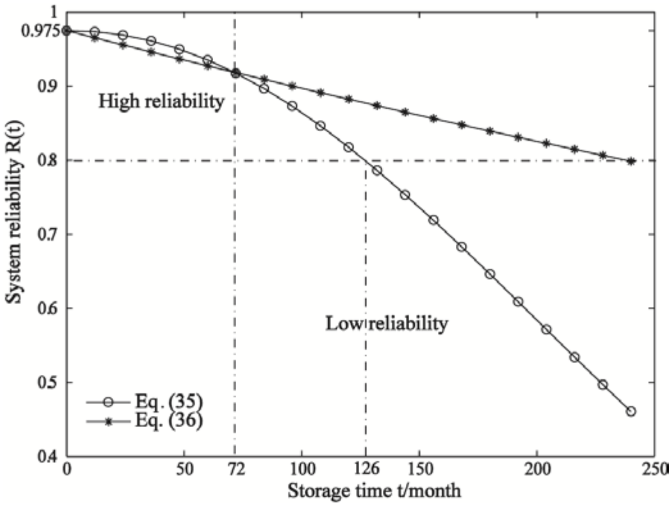


Fig. 3. Real-time reliability with complete lifetime data

From Fig. 3, it can be concluded that the degradation of Weibull type system is slower than that of exponential type, and exponential type appears more sensitive than Weibull type. However, as time goes by, exponential type exhibits higher physical strength and its mechanical behavior changes slower compared with Weibull type. Under a certain maintenance strategy, exponential type system keeps higher reliability within 20 years' storage period, while the reliability of Weibull type will be lower than 0.8 and enter into a low reliability area after 10 years.

In the next section, we choose exponential type system to verify the effectiveness of the proposed completion method and models under I-AGAN and I-ABAO.

5.2. Reliality analysis under I-AGAN

Based on the original exponential censored data (Table 1), when I-AGAN is adopted, the estimator of parameter λ_0 and β in Eq. (30) can be obtained using numerical analysis and iterative method: $\hat{\lambda}$=0.003865, $\hat{\beta}$=0.1102. Therefore, the real-time reliability under I-AGAN can be expressed as:

$$R(t) = 0.975\exp\left[-0.003865\exp\left(0.1102\cdot[\frac{t}{12}]\right)\right]\cdot\left(t-12\cdot[\frac{t}{12}]\right) \quad (37)$$

where $[\cdot]$ represents rounding down.

To validate the effectiveness of the estimation model, reliability curves drawn according to Eqs. (36), (37) are shown in Fig. 4.

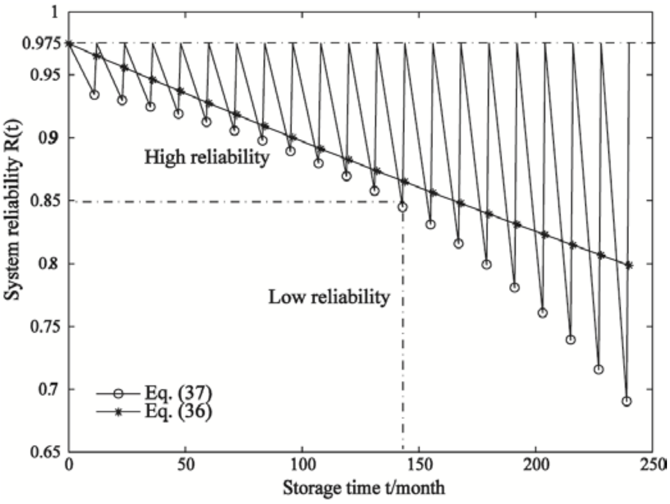


Fig. 4 Real-time reliability with censored data under I–AGAN

It can be concluded from Fig. 4 that during storage under I-AGAN, the reliability of exponential system decreases progressively faster over time, which accords with engineering practice. Within the first 10 years, the system keeps a reliability higher than 0.85, and within 19 years, the reliability still remains higher than 0.7.

5.3. Reliality analysis under I-ABAO

When I-ABAO is adopted, based on the data in Table 1 and Eq. (33), estimator of parameter λ_0, δ, β in Eq. (34) can be obtained using iterative method: $\hat{\lambda}_0$=0.001098, $\hat{\beta}$=0.2015, $\hat{\delta}$=0.000384.Therefore, the real-time reliability under I-ABAO is expressed as:

$$R(t) = 0.975\exp\left[-0.001098\exp\left(0.2015\cdot[\frac{t}{12}]\right)\cdot\left(t-12\cdot[\frac{t}{12}]\right)-0.000384t\right] \quad (38)$$

where [·] represents rounding down.

Reliability curves drawn according to Eqs. (36), (38) are shown in Fig. 5.

In Fig. 5, the reliability of exponential system under I-ABAO has similar downtrend with that under I-AGAN. Within the first 10 years of storage, the system keeps a reliability higher than 0.80. However, after 10 years, the reliability enters into a low reliability area or even fails. Compared with Fig. 4, I-AGAN can keep higher reliability.

Fig. 4 and Fig. 5 show the comparison between the reliability curves obtained from complete data and censored data under I-

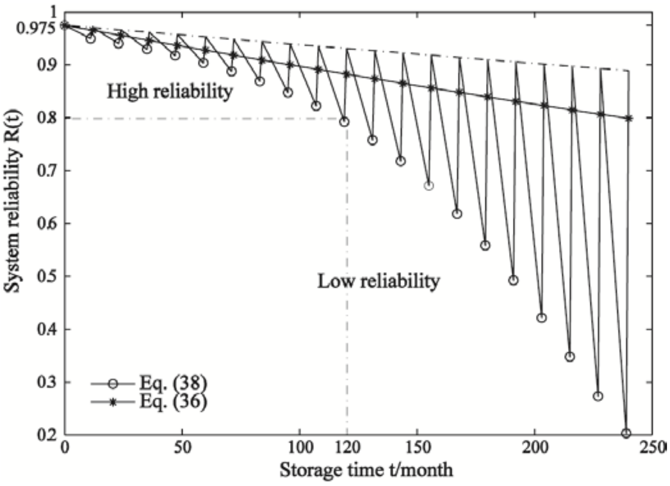


Fig. 5 Real-time reliability with censored data under I–ABAO

AGAN, I-ABAO respectively. The results show they share the similar downtrend and they are in accordance with engineering practice, which verifies the effectiveness of the proposed completion method and models.

6. Conclusion

In this paper, we develop two forecast models for real-time reliability of storage system under two repair modes. A completion method is also proposed to convert censored data into complete data by averaging the residual lifetimes. The case study shows that the system reliability estimated using the complete data shares similar changing trend with that using directly recorded censored data. The result verifies the feasibility of the proposed completion method and the effectiveness of the two models. For systems in long-time storage state under periodic inspection, the real-time reliability can be effectively estimated by applying the forecast models proposed in this paper.

In this article, it is assumed that the repair cycle is a fixed value. It should be interesting to carry out further studies on the unfixed repair cycle. Furthermore, if it is difficult to identify a suitable theoretical distribution for an application, a nonparametric approach can be employed to estimate the probability distribution based on the periodic inspection data.

***Acknowledgement**:*

This project is supported by National Natural Science Foundation of China (Grant No. 50405021), the Key Program of Science Foundation of Higher Education Institutions of Anhui Province, China (Grant No. KJ2010A337).

References

1. Anderson W T, Christianson K A, Monglestue C. Failure mechanisms in life-tested HEMTs. Quality and Reliability Engineering International 1993; 9(4): 367-370.
2. Aneziris O N, Papazoglou I A, Lygerou V. Dynamic safety analysis of process systems with an application to a cryogenic ammonia storage tank. Journal of Loss Prevention in the Process Industries 2000; 13(2): 153-165.
3. Cao R, Chen Y, Kang R. Storage reliability evaluation of engine control circuit module. Bangkok Thailand, 2011; IEEE International Conference on Quality and Reliability.
4. Cui L R, Zhao X, Shen J B, Xu Y. An availability model for storage products under periodical inspections. International Journal of Reliability, Quality and Safety Engineering 2010; 17(2): 89-103.
5. Han L, Ma Y H, Hong J, Dou M S. Research on Bayesian approach for storage reliability assessment of missile-engine. Guiyang China, 2011; 9th International Conference on Reliability, Maintainability and Safety.
6. Kodo I, Toshio N. An optimal inspection policy for a storage system with high reliability. Microelectronics Reliability 1995; 36(6): 875-882.
7. Laurent D, Oliver G. Modeling and assessment of aging and efficiency of corrective and planned preventive maintenance. IEEE Transactions on Reliability 2011; 60(4): 759-769.

8. Merren G T. Dormant storage reliability assessments-data based. IEEE Transaction on Components, Hybrids, and Manufacturing Technology 1981; 4(4): 446-454.
9. Peng W, Huang H Z, Zhang X, Liu Y, Li Y. Reliability based optimal preventive maintenance policy of series-parallel systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; (2): 4-7.
10. Timothy S. Numerical analysis. Boston : Pearson Addison Wesley, 2006.
11. Yang Y, Fu G C, Wan B, Gu H T. Component storage reliability evaluation method based on multi-performance parameters degradation analysis. Shenzhen China, 2011; Prognostics and System Health Management Conference.
12. Zhao J M, Jia X S. An optimal policy of inspection for a component with delayed repair. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; (3): 20-23.
13. Zhao M and Xie M. A model of storage reliability with possible initial failures. Reliability Engineering and Systems Safety 1994; 43(3): 269-273.
14. Zhang Y J, Zhao M. Storage reliability model and parameter estimate based on periodic inspection. System Engineering Theory and Practice 2008; 28(10): 82-88.

Associate Professor Chun Su, Ph.D.
Department of Industrial Engineering
School of Mechanical Engineering, Southeast University
Jiangning District, Nanjing 211189, Jiangsu Province, China
E-mail: suchun@seu.edu.cn

Mr. Yong-jin ZHANG
School of Mathematics and Physics
Anhui University of Technology
Ma'anshan 243002, Anhui Province, China
E-mail: zyjgltj@163.com

Ms. Bai-xue CAO
Department of Industrial Engineering
School of Mechanical Engineering, Southeast University
Jiangning District, Nanjing 211189, Jiangsu Province, China
E-mail: greenysun@163.com