POLISH
ACADEMY
OF SCIENCES

COMMITTEE
OF MACHINE
ENGINEERING

SCIENTIFIC PROBLEMS OF MACHINES OPERATION AND MAINTENANCE

ZAGADNIENIA EKSPLOATACJI MASZYN

TRIBOLOGY • RELIABILITY • TEROTECHNOLOGY
DIAGNOSTICS • SAFETY • ECO-ENGINEERING

TRIBOLOGIA • NIEZAWODNOŚĆ • EKSPLOATYKA
DIAGNOSTYKA • BEZPIECZEŃSTWO • EKOINŻYNIERIA

3 (163)
Vol. 45
2010

Institute for Sustainable Technologies – National Research Institute, Radom

EDITORIAL BOARD:

Editor in Chief	Stanisław Pytko
Deputy Editor in Chief	Marian Szczerek
Editor of Tribology	Marian Szczerek
Editor of Reliability	Janusz Szpytko
Editor of Terotechnology	Tomasz Nowakowski
Editor of Diagnostics	Wojciech Moczulski
Editor of Safety	Kazimierz Kosmowski
Editor of Eco-Engineering	Zbigniew Kłos
Scientific Secretary	Jan Szybka
Secretary	Ewa Szczepanik

EDITORIAL ADVISORY BOARD

Alfred Brandowski, Tadeusz Burakowski, Czesław Cempel, Wojciech Cholewa, Zbigniew Dąbrowski, Jerzy Jaźwiński, Jan Kiciński, Ryszard Marczak, Adam Mazurkiewicz, Leszek Powierża, Tadeusz Szopa, Wiesław Zwierzycki, Bogdan Żółtowski
and
Michael J. Furey (USA), Anatolij Ryzhkin (Russia), Zhu Sheng (China), Gwidon Stachowiak (Australia), Vladas Vekteris (Lithuania).

Mailing address: Scientific Problems of Machines Operation and Maintenance
Institute for Sustainable Technologies – National Research Institute
ul. Pułaskiego 6/10, 26-600 Radom, Poland
Phone (48-48) 364 47 90
E-mail: ewa.szczepanik@itee.radom.pl

The figures have been directly reproduced from the originals submitted by the Authors.

Publishing House of Institute for Sustainable Technologies – National Research Institute
26-600 Radom, K. Pułaskiego 6/10 St., phone (48-48) 364-42-41, fax (048) 364-47-65
www.tribologia.org

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

CONTENTS

Z. Rymuza, S. Pytko: Scaling in friction experiments 7

M.W. Sułek, M. Zięba, T. Wasilewski, W. Sas: The role of silicone derivatives as additives modifying lubricity of water 19

M. Młyńczak: Failure models of mechanical objects 29

T. Nowakowski, S. Werbińska-Wojciechowska: Maintenance models – role of logistic support review 45

J. Szybka: Methodology for reliability estimation of systems with sliding reserve 65

J. Ziółkowski: Method of formulating the required number of tankers for delivery aircrafts in aviation fuel 75

S. Duer: Examination of the effectiveness of the operation process of a radar system 87

A. Katunin: B-spline wavelet packets and their application in the multiresolution non-stationary signal processing 103

SPIS TREŚCI

Z. Rymuza, S. Pytko: Problemy skali w badaniach tarcia 7

M.W. Sułek, M. Zięba, T. Wasilewski, W. Sas: Rola pochodnych silikonów jako dodatków modyfikujących właściwości smarne wody 19

M. Młyńczak: Modele uszkodzeń obiektów mechanicznych 29

T. Nowakowski, S. Werbińska-Wojciechowska: Modele utrzymania systemu technicznego – rola wsparcia logistycznego, przegląd 45

J. Szybka: Metodyka szacowania niezawodności układów z przesuwającą się rezerwą 65

J. Ziółkowski: Metoda obliczania niezbędnej liczby cystern do zaopatrywania statków powietrznych w paliwo lotnicze 75

S. Duer: Badanie efektywności procesu eksploatacji zestawu stacji radiolokacyjnej 87

A. Katunin: B-splajnowe pakiety falkowe i ich zastosowanie w wielorozdzielczej analizie sygnałów niestacjonarnych 103

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

Z. RYMUZA*, S. PYTKO**

Scaling in friction experiments[1]

Key words

Scaling of friction, micro/nanoscale friction, adhesion.

Słowa kluczowe

Skalowanie przy tarciu, tarcie w mikro/nanoskali, adhezja.

Summary

The tribological problems, which have to be taken into consideration in micromachines, micromechanisms, Micro Electro Mechanical Systems (MEMS) devices and components such as silicon micromotors, toothed gears, gas turbines, and others are discussed. In such systems entirely different problems appear as compared with tribological problems in macroscale machines. The research proved that the values of friction parameters in the devices where the contact area is some millimeters square cannot be applied to the situation when the realistic contact area is only some micro- or nanometers square as well when the geometrical dimensions are changed.

1. Introduction

In this paper we are going to discuss problems of dry friction and the effect of scale on the results obtained in the measurement of friction. It means that we

* Department of Mechatronics (Institute of Micromechanics and Photonics), Warsaw University of Technology, Warsaw, Poland.
** Department of Machine Design and Exploitation of the Science and Technology University (AGH), Cracow, Poland.

[1] Paper presented during the First International Brazilian Conference on Tribology TriboBr, 24th to 26th,210, Rio de Janeiro, RJ, Brazil.

are going to analyse the friction without any third material (body) intentionally introduced between those two rubbing solids. In the technical literature, the friction coefficient is treated as the measure of motion resistance and its value for a number of tribological combinations of materials is given in handbooks. Many frequently published values of friction coefficients are proposed without a precise description of the research background and the operational conditions of rubbing elements The use in the design process of these values is dissatisfying both when applied to design tribosystems embodied in large machines or miniature mechanisms. One of the authors has recognised this problem when the author tried to select the coefficient of adhesion between a wheel and a rail [1]. After tests, it appeared that the adhesion coefficients given in handbooks are much higher than those observed during the investigations carried out in the laboratory with the use of real steel elements.

Very difficult problems connected with high friction are encountered in MEMS devices and magnetic recording systems. MEMS microdevices are fabricated using silicon planar technology, LIGA, or other special techniques of manufacturing [2–5]. The frictional interactions between contacting surfaces in such systems result from very strong adhesive bonds that are caused by the activity of surface molecular forces. If the volume of a component decreases, the ratio of the surface to volume increases, so that the surface interactions dominate the frictional process. A large lateral force required to initiate relative motion between two smooth surfaces is referred to as “stiction,” which has been studied extensively in the tribology of magnetic storage systems [6]. Friction/stiction (static friction), wear, and surface contamination affect device performance and, in some cases, can prevent a device from working.

The differences in geometry and size of the practical tribosystems needs intensive studies to find optimum models in experimental studies of the friction and wear behaviour of tribosystems. In every tribological test it is essential to assume the tribological model that should form the adequate representation of a real system. The scale of the model used in experiments and test conditions seriously influences the applicability of the results obtained for the prediction of the tribological behaviour of a real system.

Research has been carried out in Europe with the objective of finding methods of the determination of the friction coefficient [7]. The aim was to compare the test results from different laboratories under rather limited test conditions accepted by the group of laboratories (31 various institutions and in this number the Institute of Terotechnology in Radom, Poland) participating in that international project called VAMAS (Versailles Project on Advanced Materials and Standards). The test conditions were as follows:

– One type of tribological apparatus were used in tests, and the materials delivered to all laboratories were of the same cast having the same structure and hardness, and the friction between steel and the aluminium oxide Al_2O_3 samples was tested.

– The surfaces of samples had the same roughness parameters, and the ambient temperature of every test was similar (special air conditioned rooms).
– The loads applied on samples (pressures) and sliding speed were the same.

It was very surprising to note that the values of the measured friction coefficient were different. These results suggest that the friction is not a simple phenomenon and the prediction of friction is a very difficult task. Tribological engineering is a difficult field of engineering and science. In particular, it also concerns the scaling problems of friction. The rules of friction are not the same in the press shown in Fig. 1 and the silicon micromotor depicted in Fig. 2. Scaling in friction is a very characteristic behaviour of frictional contacts [9–12].

This statement will be supported by the results obtained performing a very simple tribological experiment. The size of the rubbing/contacting surfaces (together with the magnitude of the applied load) have been decreased considerably from one test to another, and the effects of these changes on the friction coefficient was observed.

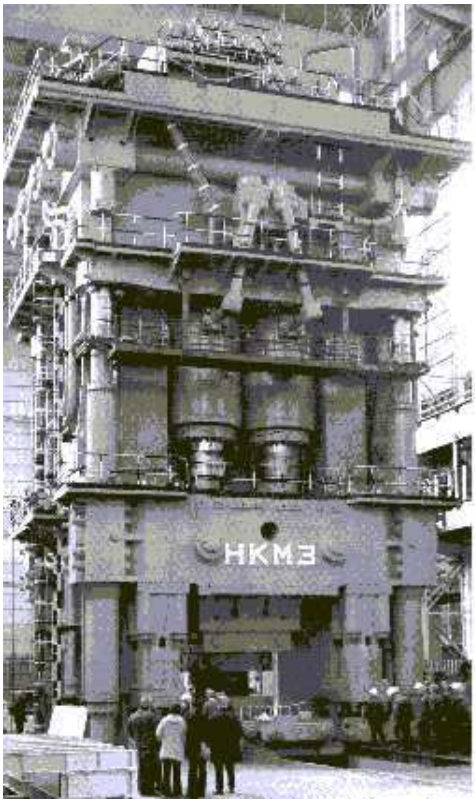


Fig. 1. Press of 650 MN manufactured in former Soviet Union and mounted in France in 1978
Rys. 1. Prasa o nacisku 650 MN wyprodukowana w dawnym ZSRR, zmontowana we Francji w 1978 roku

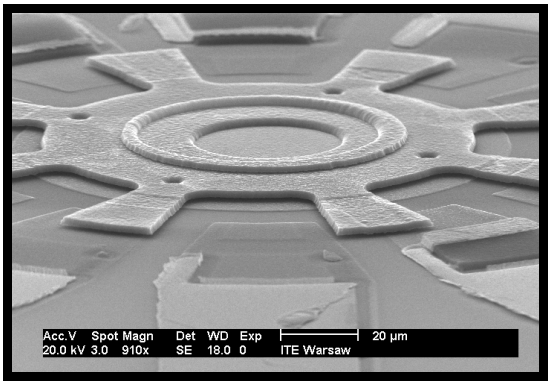


Fig. 2. Silicon micromotor manufactured in
Institute of Electron Technology in Warsaw (Poland)
Rys. 2. Silnik krzemowy wyprodukowany
w Instytucie Technologii Elektronowej w Warszawie

2. Experimental

The test rig used to carry out the friction experiments [8] is presented in Fig. 3.

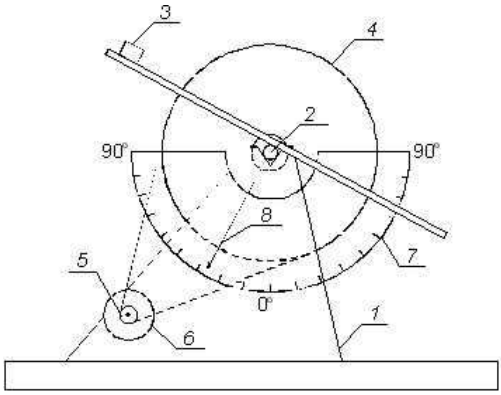


Fig. 3. Test rig used in friction experiments. 1 – base, 2 – inclined plane, 3 – sample, 4, 5, 6 – two disks and string respectively, 7 – protractor, indicator of the angle [8]
Rys. 3. Stanowisko badawcze użyte w eksperymentach tarciowych. 1 – podstawa, 2 – nachylona płyta, 3 – próbka, 4, 5, 6 – odpowiednio dwie tarcze i struna, 7 – podziałka do odczytu kąta

The simple inclined plane to measure the static friction coefficient was very useful, since the friction coefficient was estimated by the measurement of the angle of inclination of the plane to the horizontal plane. The gravity force was used to load the rubbing element.

The aluminium samples (15 μm thick foil, folded due to its large area) of the selected weight and having rather low surface roughness have been placed on

the inclined polished steel plate (Fig. 3). Prior to the test, both the sample surface and the inclined plane have been carefully cleaned using cleaning solvents and finally by the use of petroleum spirits. The experiments have been started with a gravity force 20 mN and additional weights up the total load 1.28 N. The area of contact of the foil was 4.76 x 10^4 mm^2. After each test (a few slides were conducted with one weight of the sample), the load was decreased by half by taking out the weights and finally part of the foil was cut off the folded foil. The experiments finished within the area of the foil of 6 mm^2 and at the load of 2.4 μN; therefore, the lowest load was 500,000 times smaller than the highest load. The lowest load is similar to the loads applied in tribological tests performed with the use of an Atomic Force Microscope (AFM).

The friction coefficient f was calculated as f = tan ρ (where ρ is the angle of inclination of the plane during starting of movement of the sample 3).

We also studied the effect of the size of the contacting element and the size of contact on the frictional behaviour. In this case, we used an atomic force microscope (AFM) equipped with MikroMasch cantilever NSC35 type C with a force constant 4.2 N/m. The sliding speed was 9 μm/s and the same normal load in the nN range was applied. The samples were polymeric resist ultrathin films with thicknesses of 75,100,150, 250, and 300 nm spin-coated on silicon substrate.

3. Results and discussion

The values of the inclination angle ρ needed to start the sliding of the sample down along the steel plate as a function of the load are shown in Fig. 4.

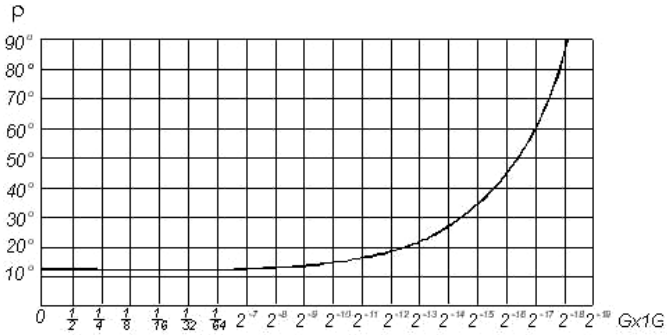


Fig. 4. Inclination angle ρ vs. applied load. 1 G = 1.28 N
Rys. 4. Kąt nachylenia ρ w funkcji obciążenia. 1 G = 1,28 N

The friction coefficient f calculated as f = tan ρ as a function of the applied load is presented in Fig. 5.

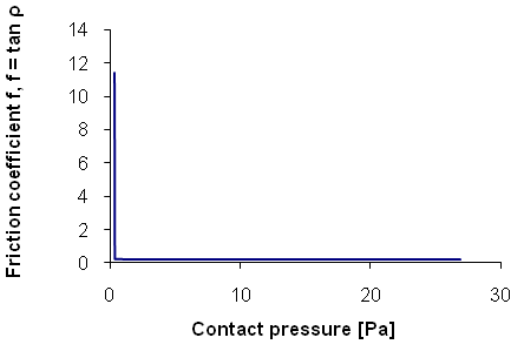


Fig. 5. Friction coefficient f = tan ρ vs applied load
Rys. 5. Współczynnik tarcia f = tg ρ w funkcji obciążenia

It is evident from this characteristic curve that at least two quite separate contributions to the friction force between these two used smooth surfaces undergoing wearless sliding: one associated with the intrinsic adhesion between two surfaces (at low loads) and the other with the externally applied load (at high loads). The "adhesion controlled" contribution to the total friction force F is proportional to the real (molecular) contact area A, and the "load controlled" friction is proportional to the load P [9]. These dependencies can be expressed as $F = \sigma A + fP$, or after dividing by the area, A as $S = F/A = \sigma + fp$; where, σ is the critical shear stress, f the coefficient of friction, P – local load, and p the local contact pressure. The coefficient of friction f is given by the slope of the friction force vs. load cure, dF/dP, rather than the absolute value of F/P; the latter is the more traditional definition of f as defined by Amonton`s law.

The friction coefficient as a function of applied averaged pressure defined as the total load divided by the contour area of contact (are of the used foil) is depicted in Fig. 6.

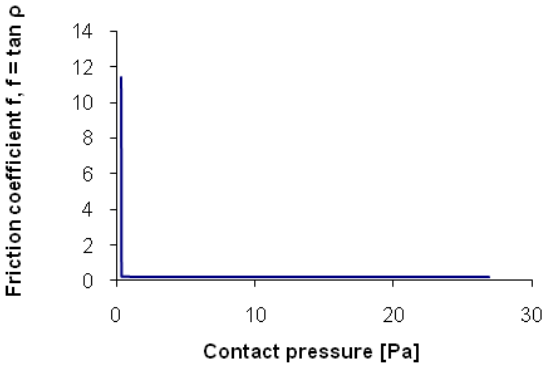


Fig. 6. Friction coefficient f = tan ρ vs. averaged contact pressure (total load divided by contour area of contact (area of aluminium foil))
Rys. 6. Współczynnik tarcia f = tg ρ w funkcji nacisku (całkowite obciążenie podzielone przez pole styku (powierzchnia folii aluminiowej))

Since the size of the rubbing element (aluminium foil) was decreased 10,000 times in the experiments, the surface-to-volume ratio was increased significantly, so the surface activity was stronger and stronger as the size (area) of the foil was decreasing. This is a characteristic situation in contacting the microcomponents in MEMS devices. The surface-to-volume ratio k was calculated from the formula $k = A_f / V = (2a^2 + 4ah)/a^2h = (2/h) + (4/a)$. A square size of the foil is assumed; a – is the square's side, h – thickness of the foil.

The curve surface-to–volume ratio k vs. the side of the square a is shown in Fig. 7.

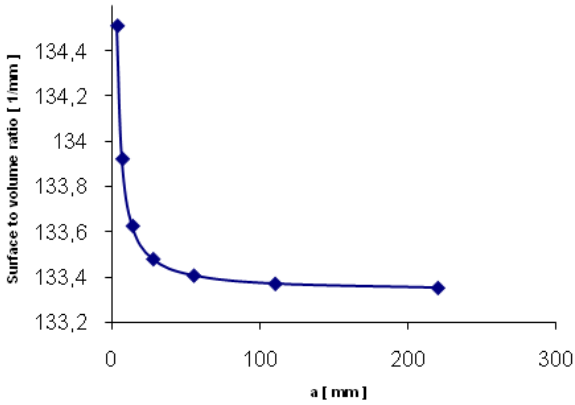


Fig. 7. Surface-to-volume ratio k vs. side of square a of aluminium foil
Rys. 7. Stosunek powierzchni do objętości k w funkcji wymiaru boku kwadratu z folii aluminiowej

The friction coefficient versus the surface-to-volume ratio k is presented in Fig. 8.

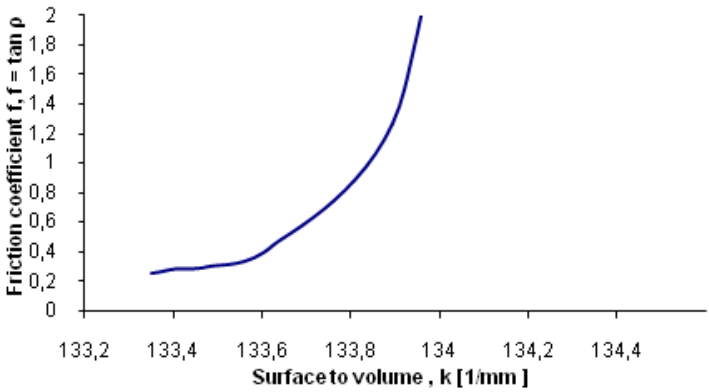


Fig. 8. Friction coefficient f = tan ρ vs. surface-to-volume ratio k
Rys. 8. Współczynnik tarcia f = tg ρ w funkcji stosunku powierzchni do objętości k

The surface activity has an important effect on the observed inclination angle and friction coefficient.

In our tribosystem constructed of a flat, polished steel plate and aluminium foil, the formation of adhesive bonds can be crucial. The tendency for two surfaces to adhere is determined by surface and interfacial energies, which are influenced by the mated materials, surface contamination, oxide layers, surface roughness, etc. [14–30]. In a broad sense, adhesion can be considered to be either physical or chemical in nature. A chemical interaction involves covalent bonds, ionic or electrostatic bonds, metallic bonds, and hydrogen bonds; and physical interaction involves the van der Waals bonds. Van der Waals forces are much weaker than in the molecules that undergo chemical interaction. These forces are always present when two asperities are in close proximity. Adhesion is a function of material pair and interface conditions such as crystal structure, crystallographic orientation, solubility of one material into another, chemical activity and separation of charges, surface cleanliness, normal load, temperature, duration of contact, and separation rate.

Let us consider a single atom strongly interacting with a rough surface displaced in a tangential direction [21] Such atoms may need to be displaced permanently during contact sliding, and such displacement of atoms can result from breakage of individual cohesive bonds or the generation of defects such as dislocations and vacancies. In the simple analysis when we neglect the effects of surface oxides and contaminants, a rough approximation for friction force can be obtained by dividing the energy required to break a cohesive bond by the distance slid, or the lattice spacing. The bond energy for the weaker material aluminium is 327 kJ/mol [22], which corresponds to 5.4×10^{-19} J/atom and a lattice spacing of 4.1×10^{-10} m, so the friction force per atom is about 1.3×10^{-9} N.

The friction force is affected by the normal load, since this force dictates the number of the atomic interactions. The prediction of the total friction force comes from the uncertainty of the number of atoms involved in the frictional interaction. We may attempt to predict the total friction force from the real area A of contact, which is typically expressed as P/H [18] (P-applied load, H – flow pressure or hardness of the softer material). For the highest applied load equal to 1.28 N and the hardness equal to about 0.3 GPa (the hardness of aluminium foil was measured on the depth about 1 μm by the nanoindentation technique by using a TriboScope® instrument of Hysitron Inc.), and the real area of contact is 4.3×10^{-8} m^2 which corresponds to a projected area of about 10^{10} atoms (the radius of A1 is 143 pm). At the load 20 mN (the full size of foil without additional weights), the real area of contact was estimated to be about 7×10^{-11} m^2. The lowest load 2.4 μN could have resulted in 8×10^{-14} m^2 approximated value of the real area of contact. The estimated number of atoms corresponding to these values of the area of contact are 1.55×10^8 and 1.8×10^5 atoms, respectively.

Using the friction force per atom 1.3 x 10^{-9} N obtained previously, the total friction force comes to about 13 N for 10^{10} atoms (applied load 1.28 N). At the applied load 20 mN and 2.4 μN, the estimated values of friction force are 200 mN and 234 μN, respectively. These values of the total friction forces cause the values of the friction coefficient to be higher than the experimental values (0.21, 0.23, and about 100, respectively). The prediction of such a high friction coefficient may have resulted from overestimating the number of atoms involved and/or the critical shear stress. The number of atoms in contact may also have little to do with the frictional force observed during sliding. It is likely that atoms with the weakest cohesive energies will be displaced during the sliding process. The crystal imperfections may cause that such energies may be orders of magnitude smaller than the ideal values. In a similar context, the number of atoms involved in the breakage of cohesive bonds should be estimated not from the contact area but from the density of such imperfections within the volume of the interacting asperities [19].

Very high friction coefficients obtained in the range of very low loads (and small area of aluminium foil) could be the effect of quite high liquid-mediated adhesive forces occurring because of the condensation of water from vapour on both contacting and near-contacting asperities. The foil was observed to be firmly stuck to the steel surface and no sliding occurred at the inclination angle 90 degrees.

In the case of the studies of ultrathin polymeric resist films, we observed the effect of the film thickness that was connected with the different size contact at the same load effected by the different deformation in the area of contact. The results of these studies are presented in Fig. 9.

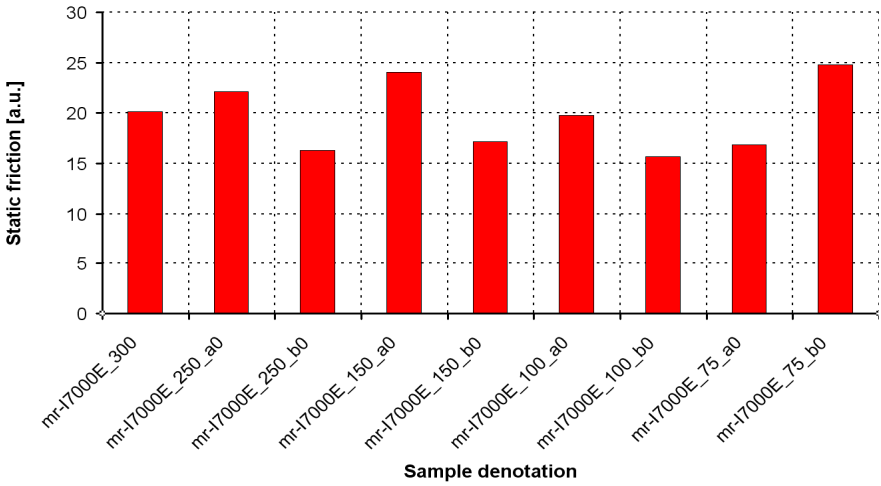


Fig. 9. Static friction vs. film thickness
Rys. 9. Współczynnik tarcia statycznego w funkcji grubości warstwy

In this case, probably the most important aspect was the effect of the deformation (mechanical) component of the friction force on observed friction. The optimum thickness of the film can be found to minimise friction in the studied contacts.

4. Conclusions

The results of the described simple experiment performed by using an aluminium foil sliding on the flat steel surface confirm the general basic equation for wearless friction describing that at low loads the friction force F is adhesion controlled ($F=\sigma A$) (σ – critical shear stress, A – real (molecular) contact area) but it is load controlled ($F=f\ P$) (f-friction coefficient, P-applied load) at high loads. The friction force is proportional to a purely load-dependent term and a purely adhesion-dependent term; the latter being proportional to the number of bonds being sheared at the junction nw (n – number of bonds broken, w – energy per bond) which may be associated with $\Delta\gamma A$ ($\Delta\gamma$ is the thermodynamic (equilibrium) surface energy or work of adhesion W_{ad}).

At low loads, strong adhesion or bonding across the interface between the aluminium foil and the flat steel surface occurred that required a finite normal force, called adhesive force, to pull the two solids apart. This effect was demonstrated by very high values of the inclination angle needed to start to slide the aluminium foil. Such an effect was observed when the values of the measured inclination angle were over 20˚, which corresponds to the situation of the applied loads being below 260 μN.

When the size of the rubbing elements decreases from 1 cm to 1 mm, the area decreases by a factor of a million and the volume decreases by a factor of a billion. As a result, surface forces such as friction, adhesion, meniscus forces, viscous drag and surface tension that are proportional to area, become a thousand times larger than the forces proportional to the volume, such as inertial or electromagnetic forces [28, 29]. The increase in resistive forces, such as friction and adhesion, because of the increase of the surface-to-volume ratio, was observed in our experiments. This is a particularly important effect in MEMS devices that are designed for small tolerances, so the physical contact becomes more likely. The high adhesion between adjacent components leads to the appearance of a large lateral force required to initiate relative motion between two smooth surfaces referred to as “stiction,” which has been studied extensively in the tribology of magnetic storage systems [6].

The effect of the thickness of the polymeric resist films in the range of thicknesses 75-300 nm was observed during the study of friction with the use of an AFM equipped with a silicon cantilever.

Acknowledgement

The partial support from NaPa European 6FP Project related to study of the resist films is acknowledged.

References

[1] Pytko S., Furmanik K., Designing track and driving wheels in cranes, Przeglad Mechaniczny No.3 (1974) 89-93, in Polish.
[2] Madou M., Fundamentals of Microfabrication, 2nd ed., CRC Press, Boca Raton 2002.
[3] Gardner J., Varadan V.K., Awadelkarim O.O., Microsensors, MEMS and Smart Devices, J.Wiley, Chichester 2001.
[4] Elwenspoek M., Wiegerink R.J., Mechanical Sensors. Springer Verlag, Berlin 2001.
[5] B. Bhushan (Editor), Tribology Issues and Opportunities in MEMS, Kluwer Academic Publishers, Dordrecht 1998.
[6] Bhushan B., Tribology and Mechanics of Magnetic Storage Devices. 2nd edition, Springer Verlag, Berlin 1996.
[7] Czichos H., Becker S., Lexow J., Multilaboratory tribotesting: Results from the Versailles Advanced Materials and Standards Programme on Wear Test Methods, Wear 114 (1987) 109-130.
[8] Krajewski S., Remarks on „dry" friction, Manuscript in the possession of S. Pytko , in Polish.
[9] Israelachvili J., Tribology of ideal and non-ideal surfaces and fluids, in: B.Bhushan Editor, Fundamentals of Tribology and Bridging the Gap Between the Macro- and Micro/Nanoscales, Kluwer Academic Publishers, Dordrecht 2001, 631-650.
[10] Zum Gahr K.H., Microtribology, Interdisciplinary Science Reviews 18 (1993) 259-266.
[11] Bhushan B., Principles and Applications of Tribology, J. Wiley, Chichester 1999.
[12] Nosonovsky M., Bhushan B., Multiscale Friction Mechanisms and Hierarchical Surfaces, Springer, Berlin 2008.
[13] Popov V.L., Contact Mechanics and Friction, Springer , Berlin 2010.
[14] Mastrangelo C.H., Adhesion-related failure mechanisms in micromechanical devices, Tribology Letters 3 (1997) 223-238.
[15] Legtenberg R., Tilmans H.A.C., Elders J., Elwenspoek M., Stiction of surface micromachined structures after rinsing and drying: model of investigation of adhesion mechanisms, Sensors and Actuators A 43 (1994) 230-238.
[16] Neuman A.W., Spelt J.K. (Editors), Applied Surface Thermodynamics, Marcel Dekker, New York 1996.
[17] Zum Gahr K.H., Microstructure and Wear of Materials, Elsevier, Amsterdam 1987.
[18] Rabinowicz E., Friction and Wear of Materials, 2nd edition, J. Wiley, Chichester 1995.
[19] Bowden F.P., Tabor D., The Friction and Lubrication of Solids, Clarendon Press, Oxford 1986.
[20] Johnson K.L., Kendall K., Roberts A.D., Surface energy and the contact of elastic solids, Proceedings of the Royal Society of London A 324 (1971) 301-313.
[21] Kim D.E., Suh N.P., On microscopic mechanisms of friction and wear, Wear 149 (1991) 199-208.
[22] Kittel C., Introduction to Solid State Physics, 7th edition, J.Wiley, Chichester 1995.
[23] Bhushan B., Introduction to Tribology, J.Wiley, Chichester 2002.
[24] Bhushan B., Contact mechanics of rough surfaces in tribology: multiple asperity contact, Tribology Letters 4 (1998) 1-35.

[25] Gao C., Tian X., Bhushan B., A meniscus model for optimization of texturing and liquid lubrication of magnetic thin-film rigid disks, Tribology Transactions 38 (1995) 201-202
[26] Komvopoulos K., Surface engineering and microtribology for microelectromechanical systems, Wear 200 (1996) 305-327.
[27] Israelachvili J., Intermolecular and Surface Forces, 2nd edition, Academic Press, San Diego 1991.
[28] Sundararajan S., Bhushan B., Micro/nanoscale tribology of MEMS materials, lubricants and devices, in: B.Bhushan Editor, Fundamentals of Tribology and Bridging the Gap Between the Macro- and Micro/Nanoscales, Kluwer Academic Publishers, Dordrecht 2001, 821-850.
[29] Rymuza Z., Control tribological and mechanical properties of MEMS surfaces, Part 1: critical review, Microsystem Technologies, 5 (1999) 173-180.
[30] Tabor D., Tribology and physics, Journal of Physics: D:Applied Physics, 25 (1992) A1-A2

Manuscript received by Editorial Board, February 2nd, 2011

Problemy skali w badaniach tarcia

S t r e s z c z e n i e

Problemy tribologiczne spotykane w mikrourządzeniach, takich jak np. MEMS i w mikro-silnikach krzemowych, a także w przekładniach zębatych, turbinach gazowych, prasach itp. różnią się znacznie między sobą. Różnice te wynikają z różnych skali wymiarowych. Stwierdzono, że zidentyfikowane tarciowe opory ruchu, gdy powierzchnia kontaktu jest w skali milimetrowej, nie mogą być wykorzystywane do prognozowania takich oporów ruchu, gdy powierzchnia kontaktu jest w skali mikro- lub nanometrowej, a także wtedy, gdy wymiary elementów trących znacznie się różnią.

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

MARIAN W. SUŁEK*, MAŁGORZATA ZIĘBA*, TOMASZ WASILEWSKI*, WITOLD SAS*

The role of silicone derivatives as additives modifying the lubricity of water

Key words

Lubricants, surfactants, cationic and nonionic silicones, physicochemical and tribological properties.

Słowa kluczowe

Substancje smarowe, związki powierzchniowo czynne, kationowe i niejonowe pochodne silikonów, właściwości fizykochemiczne i tribologiczne.

Summary

Silicone oils are used as lubricants or additives modifying lubricating properties of oil bases. However, those compounds are not water soluble. Therefore, the following silicone derivatives were selected as additives improving the lubricity of water: quaternary diamide polydimethylsiloxane (Quaternium-80) which is a cationic surfactant and an ethoxylated silicone which is a nonionic surfactant (Bis-PEG/PPG-20/20 Dimethicone). Surface activity of the two compounds was high and it was confirmed by surface tension measurements. In this study, motion resistance and wear of steel were determined in the presence of aqueous solutions of selected silicone derivatives as model lubricating substances. The tests were carried out using a T-11 testing apparatus produced by the Institute for Sustainable Technologies in Radom. It has been found that the cationic surfactant considerably decreases motion resistance in comparison with its nonionic equivalent in the low concentration range of the order of a few percent.

* Technical University of Radom, Department of Chemistry, ul. Chrobrego 27, 26-600 Radom.

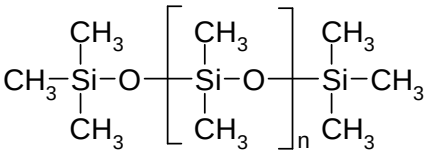
1. Introduction

Lubricants based on silicone oils have been used in lubrication practice for a number of years [1–3]. They exhibit high thermal resistance, low susceptibility to oxidation and are resistant to the action of aggressive chemical agents. They can be used in a wide temperature range of from sub-zero temperatures of the order of about fifty degrees to above-zero temperatures of the order of several hundred degrees Celsius. Silicones are physiologically and biologically neutral. This brief characterisation indicates that the presence of a silicone chain in a molecule makes it possible to formulate silicone-based materials having properties that cannot be obtained using hydrocarbons [1-3].

The object of this investigation are lubricating substances based on water whose poor lubricating properties are modified by surface active compounds used as additives [4-6]. Quaternary diamide polydimethylsiloxane (INCI: Quaternium-80, trade name: Abil Quat 3474, manufacturer: Evonik Chemicals) with positive charges at two quaternary nitrogen atoms was selected for this investigation. The results of physicochemical and tribological tests obtained for aqueous solutions of Quaternium-80 (cationic surfactant) were referred to in an analogous study carried out for a linear silicone polyether (INCI: Bis-PEG/PPG-20/20 Dimethicone trade name: Abil 8832, manufacturer: Evonik Chemicals). The hydrophilic part of this nonionic surfactant was an ethylene oxide chain while a silicone backbone formed the hydrophobic part.

2. Properties of silicones and silicone polyethers

Silicone oils (Formula 1) are clear, tasteless, odourless liquids. They exhibit high resistance to low temperatures (up to -50^0 C) and high temperatures (up to 250^0 C). Their viscosity (from several to as much as a few million mPa·s) depends on their molecular mass and varies slightly with temperature. Silicones are resistant to oxidation and to the action of aqueous solutions of acids, bases, and salts. They do not hydrolyse readily [1, 7, 8].



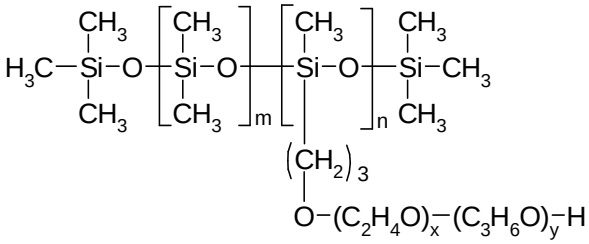
(1)

Structural formula of polydimethylsiloxane
Wzór strukturalny polidimetylosiloksanu

Silicone oils exhibit good wettability of various types of materials. They have considerable hydrophobing properties and are used as antifoaming agents. The compounds are physiologically neutral and safe for the environment [1, 7, 8].

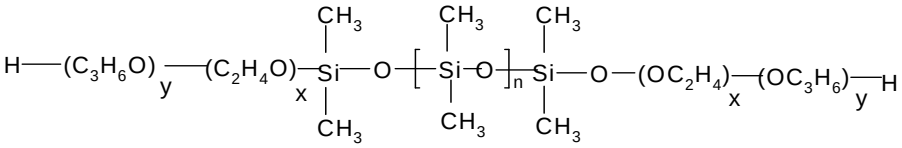
Lubricant films can break and be squeezed out of the friction zone by external loads under friction conditions. This is more likely in the case of a metal/metal contact, but less likely in the case of metal/plastic or plastic/plastic contacts [2].

An attempt has been made in this study to apply silicone derivatives that, apart from the hydrophobic part, would have a hydrophilic part [9]. These derivatives retain the properties of silicones but, additionally, show affinity for water and hydrophilic surfaces [7–8]. The compounds having such properties are silicone glycols, also called silicone polyethers, polyether polysiloxanes or PEG/PPG Dimethicone (Formulas 2 and 3).



(2)

Structural formula of comb-like silicone glycol
Wzór strukturalny glikolu silikonowego o budowie grzebieniowej



(3)

Structural formula of block silicone glycol
Wzór strukturalny glikolu silikonowego o budowie blokowej

Silicone polyethers are a group of nonionic surfactants. They are capable of forming micelles in solutions but, in contradistinction to hydrogen-based surfactants, their surface activity is higher [7-8].

Surface activity, solubility, and miscibility with solvents and other constituents can be modified at the stage of the synthesis of the compounds. In order to obtain the desired properties, the compounds are assigned an appropriate structure, an organosilicone chain length, and the lengths of ethylene oxide (PEG) and propylene oxide (PPG) chains. An increase in the number of PEG groups leads to an increase in water solubility, and an increased number of PPG groups increase solubility in nonpolar media [1].

The application of silicone glycols as additives modifying lubricating properties of water is interesting [9]. Aqueous solutions of silicone glycols (c = 1 - 80%) and "pure" compounds (c = 100%) were suggested as model lubricants. The effectiveness of lubricating substances was tested based on the estimation of motion resistance and antiwear properties.

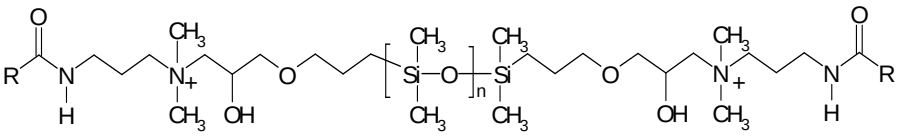
Based on the tests carried out at constant loads (10 and 50 N) and using a ball-on-disc tribosystem (T-11 testing apparatus), it was shown that there was no decrease in motion resistance and wear at low concentrations of polyethers (1, 4, 10%) [9]. It can even be stated that the quantities tend to increase slightly relative to water. Friction conditions change dramatically in the case of the concentrations of the order of forty percent, where the value of μ drops as much as nearly fourfold. The character of μ(c) changes at both 10 N and 50 N loads is similar. The lowest μ values (as low as 0.07) were observed for pure compounds at both 10 N and 50 N. Such low motion resistance indicate fluid friction in which friction couple surfaces are fully separated by a lubricant film. The results of the tests conducted at 10 and 50 N showed that the wear of friction couples decreased with concentration of the compounds. The lowest wear was observed in the case of pure compounds (wear scar diameters were even more than threefold smaller than those for water) [9].

3. Structure and surface activity of selected silicone derivatives

As mentioned above, silicone polyethers (nonionic surfactants) efficiently modify lubricity of water at the concentrations of forty percent. However, at the concentrations of several percent, the effect of the additive is modest and it is even possible to observe an increase in motion resistance and wear [9].

In view of the beneficial properties of silicones and their derivatives, as well as the concentration range in which the additives are used (of the order of 1%), an attempt was made to search for other silicone derivatives that would be effective at low concentrations.

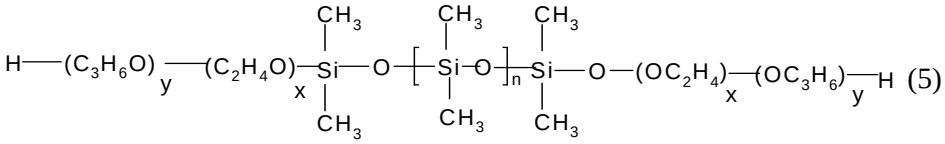
The analysis of literature data and our own research indicates that the formation of the surface phase is connected not only with the hydrophobic effect but also with the interactions of molecules with the solid surface [4, 10]. In the case of nonionic surfactants, these are hydrogen bonds and also weak dispersive interactions. They may not be strong enough to produce the surface phase that, under friction, transforms into a lubricant film having high resistance to mechanical and thermal loads. Therefore, an ionic surfactant (a cationic organosilicone compound) was selected for the studies as an additive to water. Its structure is presented below (see Formula 4).



(4)

Structural formula of Quaternium-80
Wzór strukturalny Quaternium-80

There are two positively charged quaternary nitrogen atoms at the ends of the silicone chain. The results obtained for the solutions of this compound will be referred to as solutions of silicone glycol (INCI: Bis-PEG/PPG-20/20 Dimethicone) that, just like Quaternium-80, is linear and having ethylene oxide and propylene oxide chains at the end of the silicone chain (Formula 5).



Structural formula of Bis-PEG/PPG-20/20 Dimethicone
Wzór strukturalny Bis-PEG/PPG-20/20 Dimethicone

Surface activity of the solutions of both compounds (Formulas 4 and 5) was compared. Its measure was surface tension (σ) determined by means of the ring detachment method (Lauda Tensiometer TD1) (Fig. 1). The measurements were made for aqueous solutions of the compounds with the concentrations of 0.001-5% - Quaternium-80 and 0.01-10% - Bis-PEG/PPG-20/20 Dimethicone.

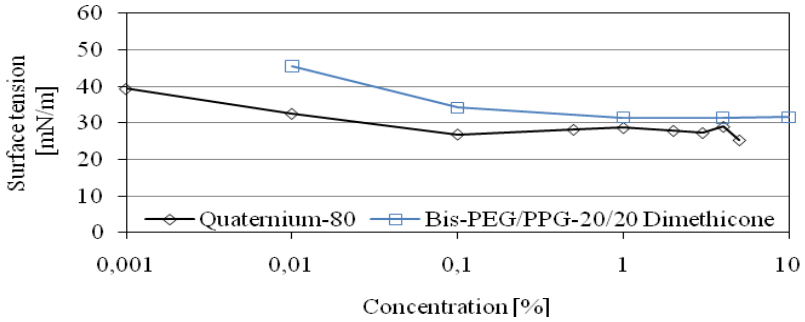


Fig. 1. Dependence of surface tension on concentration of aqueous solutions of Quaternium-80 and Bis-PEG/PPG-20/20 Dimethicone. The value of surface tension for water is about 72 mN/m
Rys. 1 Zależność napięcia powierzchniowego od stężenia wodnych roztworów Quaternium-80 oraz Bis-PEG/PPG-20/20 Dimethicone. Wartość napięcia powierzchniowego wody ok. 72 mN/m

The differences found can be interpreted as a result of stronger interactions of the ionic surfactants with the surface.

Viscosity is one of the more important quantities that may affect tribological properties of a lubricating substance. Therefore, viscosity of the solutions as a function of additive concentration was compared just before the tribological tests. The studies were carried out using aqueous solutions of silicones of the same concentrations as those used in surface tension measurements. The results are given in Fig. 2.

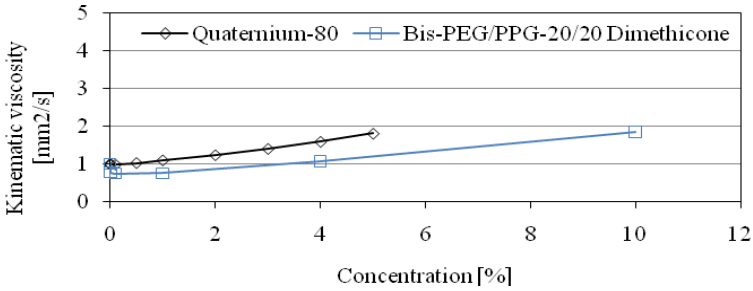


Fig. 2. Dependence of kinematic viscosity on concentration of aqueous solutions of Quaternium-80 and Bis-PEG/PPG-20/20 Dimethicone
Rys. 2. Zależność lepkości kinematycznej od stężenia wodnych roztworów Quaternium-80 oraz Bis-PEG/PPG-20/20 Dimethicone

The values of viscosity coefficients are so low in the concentration ranges used that they will not affect tribological properties of the lubricants tested.

4. Tribological properties

The evaluation of tribological properties was carried out for 1-5% aqueous solutions of Quaternium-80 and 1–10% aqueous solutions of Bis-PEG/PPG-20/20 Dimethicone. The measurements were made for 3 loads (10, 30, and 50 N) to determine the coefficient of friction (μ) and wear scar diameter (d). The methods presented in the literature [11] were applied. For tribological measurements steel samples (ŁH15) were used. Ball of the diameter equal to 0.5" was used as a sample. The disc of the diameter equal to 25 mm, surface roughness Ra = 0.16 μm was used as a counter-sample.

The solutions of cationic surfactants have lower friction coefficients than water, which is the base of the model lubricants analysed (Fig. 3).

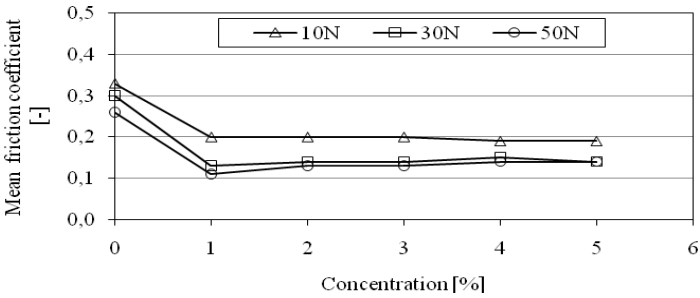


Fig. 3. Dependence of mean friction coefficient on concentration of aqueous solutions of Quaternium-80. Tribological test, T-11 testing apparatus, steel ball-steel disc couple, test duration 900 s, sliding velocity 0.1 m/s, radius 10 mm, loads 10, 30 and 50 N
Rys. 3. Zależność średniego współczynnika tarcia od stężenia wodnych roztworów Quaternium-80. Badanie tribologiczne, aparat T-11, skojarzenie kulka stalowa–tarcza stalowa, czas testu 900 s, prędkość poślizgu 0,1 m/s, promień 10 mm, obciążenia 10, 30 oraz 50 N

After using Quaternium-80 in the lowest concentration (c = 1%) μ values were even 2 times lower in comparison to water, as a base. Further growth of cationic silicone concentration generally didn`t cause changes in the μ values. The reason of changes could be related with Quaternium-80 surface activity (Fig. 1) and the possibility of micells appearance in water solutions.

As it has already been mentioned in Chapter 2, motion resistance for a nonionic surfactant (Bis-PEG/PPG-20/20 Dimethicone) are relatively high, comparable and even higher than those observed in the presence of water (Fig. 4).

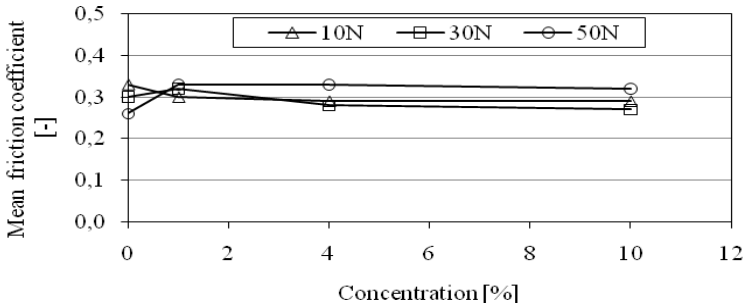


Fig. 4. Dependence of mean friction coefficient on concentration of aqueous solutions of Bis-PEG/PPG-20/20 Dimethicone. Tribological test, T-11 testing apparatus, steel ball-steel disc couple, test duration 900 s, sliding velocity 0.1 m/s, radius 10 mm, loads 10, 30 and 50 N
Rys. 4. Zależność średniego współczynnika tarcia od stężenia wodnych roztworów Bis-PEG/PPG-20/20 Dimethicone. Badanie tribologiczne, aparat T-11, skojarzenie kulka stalowa–tarcza stalowa, czas testu 900 s, prędkość poślizgu 0,1 m/s, promień 10 mm, obciążenia 10, 30 oraz 50 N

In the case of all the concentrations and loads, the μ values are higher than threefold for Bis-PEG/PPG-20/20 Dimethicone solutions than for Quaternium-80 (Figs 3 and 4). The largest changes in the μ value as a function of Quaternium-80 concentration can be observed up to 1%. Above this concentration, the coefficient of friction stabilises within the measurement error limits (0.01). Another tendency observed is a drop in the μ value with an increase in load. The drop is quite considerable with the load change from 10 to 30N but less so with the 30 to 50N change (Fig. 3). The maximum decrease in the μ value for individual loads (10, 30, 50N) relative to water is 1.6, 2.3, and 2.4, respectively. Thus, at moderately high loads (10-50N) the lowest decrease was observed for 10N, while the decreases for 30 and 50N are comparable and higher than those for 10N.

The analysis of Fig. 5 indicates that the action efficiency of Quaternium-80 as an antiwear additive is low.

Wear scar diameters (d) obtained after applying Quaternium-80 range from 0.58 to 0.70 mm (see Fig. 5), and thus they are larger for the model lubricants proposed than for water (d = 0.50 - 0.57). The results may indicate chemical wear resulting from the reaction of ionic surfactants with the steel surface.

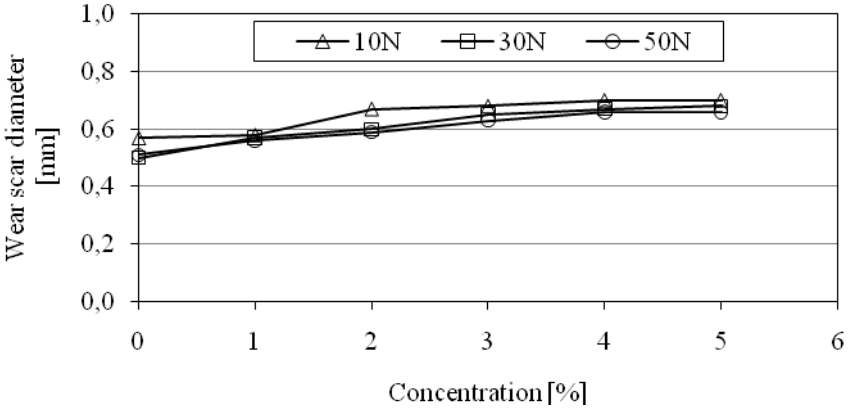


Fig. 5. Dependence of wear scar diameter on concentration of aqueous solutions of Quaternium-80. Tribological test, T-11 testing apparatus, steel ball-steel disc couple, test duration 900 s, sliding velocity 0.1 m/s, radius 10 mm, loads 10, 30 and 50 N

Rys. 5. Zależność średnicy skazy kulki od stężenia wodnych roztworów Quaternium-80. Badanie tribologiczne, aparat T-11, skojarzenie kulka stalowa–tarcza stalowa, czas testu 900 s, prędkość poślizgu 0,1 m/s, promień 10 mm, obciążenia 10, 30 oraz 50 N

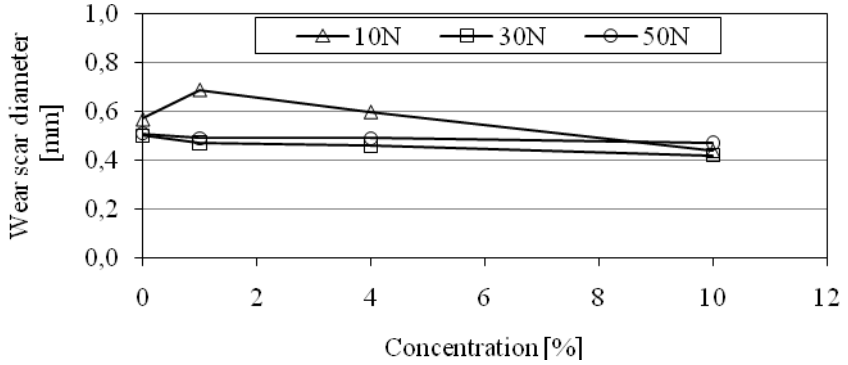


Fig. 6. Dependence of wear scar diameter on concentration of aqueous solutions of Bis-PEG/PPG-20/20 Dimethicone. Tribological test, T-11 testing apparatus, steel ball - steel disc couple, test duration 900 s, sliding velocity 0.1 m/s, radius 10 mm, loads 10, 30 and 50 N

Rys. 6. Zależność średnicy skazy kulki od stężenia wodnych roztworów Bis-PEG/PPG-20/20 Dimethicone. Badanie tribologiczne, aparat T-11, skojarzenie kulka stalowa–tarcza stalowa, czas testu 900 s, prędkość poślizgu 0,1 m/s, promień 10 mm, obciążenia 10, 30 oraz 5 0N

Nonionic surfactants exhibit slightly better antiwear properties (Fig. 6). Wear of the balls decreases slightly as a function of increasing concentration of an additive relative to wear observed when water is used as a base. The maximum drop in the d value of the solutions relative to water is several percent (4% solutions).

5. Summary

Organosilicone compounds possess a number of interesting physicochemical properties that form the basis for their application as active constituents of lubricating substances. Silicones are, however, insoluble in water and that considerably limits their application range. Therefore, the subjects of this study are water-soluble silicone polymers [9]. Used as additives, they show high efficiency only in the case of high concentrations of the order of forty percent. As additives modifying tribological properties of bases, they are used primarily at the concentrations of the order of several percent, and an attempt was made to find silicone derivatives which would improve tribological properties in this concentration range.

The silicone selected for the investigation was an ionic one (Quaternium-80) and the investigation results were referred to water as a base and to aqueous solutions of a nonionic silicone (Bis-PEG/PPG-20/20 Dimethicone). The analysis of the results obtained indicates that, for the ball-disc tribosystem (T-11 testing apparatus) and loads of above 10 N, the ionic surfactant considerably affects a reduction in the coefficient of friction both relative to water (about twofold) and relative to the nonionic silicone solutions (about threefold). Wear, however, is comparable or even higher than the one for water.

The beneficial effect of the ionic silicone additive on the reduction in motion resistance can be explained by stronger, compared to the nonionic surfactant, interactions with the steel surface; whereas, high wear can be attributed to the chemisorption of additives.

The application of silicone water solutions as lubricants is predicted in friction pairs were leakage or human skin contact may appear, and in machines used in cosmetic industry, pharmaceutic and household products manufacturing.

Acknowledgments

The research was financed by Ministry of Science and Higher Education in the years 2007–2010 – Project No. N 507 111 32/3345.

References

[1] Rościszewski P., Zielecka M.: Silikony – właściwości i zastosowanie, Wydawnictwo Naukowo-Techniczne, Warszawa, 2002.

[2] Capporiccio G., Cann P.M., Spikes H.A.: Additives for fluorosilicone oils at high temperature, Wear, 193, 1996, 261-268.

[3] Ni S.C., Kuo P.L., Lin J.F.: Antiwear performance of polysiloxane-containing copolymers at oil/metal interface under extreme pressure, Wear, 253, 2002, 862-868.

[4] Sułek M.W.: Wodne roztwory surfaktantów w inżynierii materiałowej systemów tribologicznych, Wydawnictwo Instytutu Technologii Eksploatacji, Radom, 2009.

[5] Sułek M.W, Wasilewski T., Influence of critical micelle concentration (CMC) on tribological properties of aqueous solutions of alkyl polyglucosides, Tribology Transaction, 52, 2009, 12-20.
[6] Sułek M.W., Wasilewski T., Zięba M.: Tribological and physical-chemical properties aqueous solutions of cationic surfactants, Industrial Lubrication and Tribology, 2010, 62/5, 279-284.
[7] Hill R.: Silicone surfactants, Marcel Dekker, New York, 1999.
[8] Maciejewski H., Marciniec B.: Polietery silikonowe. Synteza, właściwości, zastosowania, Przemysł Chemiczny, 2(82), 2006, 138-141.
[9] Sułek M.W., Zięba M., Polietery silikonowe jako efektywne dodatki modyfikujące właściwości smarne wody, Tribologia, 3/2009, 226, 247-256.
[10] Nagatani N., Fukuda K., Suzuki T.: Interfacial behaviour of mixed systems of glycerylether-modified silicone and polyoxyethylene-modyfied silicone. Journal of Colloid and Interface Science, 234, 2001, 337-343.
[11] Piekoszewski W., Szczerek M., Wulczyński J: Urządzenie typu trzpień-tarcza do badań tribologicznych w podwyższonych temperaturach, Tribologia, 5-6, 1997, 826-832.

Manuscript received by Editorial Board, January 10th, 2011

Rola pochodnych silikonów jako dodatków modyfikujących właściwości smarne wody

S t r e s z c z e n i e

Oleje silikonowe są stosowane jako substancje smarowe względnie dodatki modyfikujące właściwości smarowe baz olejowych. Związki te wykazują jednak brak rozpuszczalności w wodzie. Dlatego jako dodatki poprawiające smarność wody wytypowano rozpuszczalne w wodzie pochodne silikonowe. Są nimi: czwartorzędowy diamidopolidimetylosiloksan (Quaternium-80), zaliczany do surfaktantów kationowych oraz oksyetylenowany silikon, należący do grupy niejonowych związków powierzchniowo czynnych (Bis-PEG/PPG-20/20 Dimethicone). Aktywność powierzchniowa obydwu związków w wodzie jest wysoka i została potwierdzona przez pomiar napięcia powierzchniowego. W prezentowanej pracy wyznaczone zostały opory ruchu i zużycie stali w obecności wodnych roztworów wybranych pochodnych silikonowych, jako modelowych substancji smarowych. Testy prowadzono za pomocą aparatu T-11, wyprodukowanego przez Instytut Technologii Eksploatacji w Radomiu. Stwierdzono, że w zakresie niskich stężeń, rzędu kilku procent, kationowy surfaktant zdecydowanie zmniejsza opory ruchu w stosunku do jego niejonowego odpowiednika.

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

MAREK MŁYŃCZAK*

Failure models of mechanical objects

Key words

Physical degradation, mechanical object, reliability, failure.

Słowa kluczowe

Fizyczna degradacja, obiekt mechaniczny, niezawodność, uszkodzenie.

Summary

Mechanical objects are operated in the real world where the degradation of material components and the variability of processes managed by man are the main factors influencing its efficiency. Machine degradation is a long-lasting process concerning its material structure, components, and connections. The loss of machine operation is due to failures caused by wear and tear, fatigue, corrosion, overloading, material ageing, and many other destructive processes. A close relation between failure modes and reliability models is observed so that knowledge about failures may help analysts create reliability models and determine the best operational decisions. This paper discusses the relation between physical phenomenon and theoretical models as a common platform of decision processes.

1. Machine and device- specifics of mechanical objects

Machines and mechanical devices are defined as technical objects usually consisting of movable elements using energy and information to process or transform energy in order to perform work on mechanical principles [1, 2]. Machines are systems of solid, usually metallic, links (bars) connected to two or

* Wroclaw University of Technology, Poland, marek.mlynczak@pwr.wroc.pl.

more other links by pin joints (hinges), sliding joints, or ball-and-socket joints, to form a closed chain or a series of closed chains [1]. The main advantage of most machines is that it multiplies human efficiency, for instance, a driver moving tons of goods with high speed, or an aeroplane or ship transporting people or goods. Machines fulfil different tasks with high efficiency, precision or at lower risk, assuring comfort and safety.

1.1. Operational system of mechanical object

Machine operation requires that the operational system and process are defined [2]. The technical object performs its function with the support of mankind (crew, staff, operators, mechanics, and managers), in a properly prepared environment: territories, operational base, task or supply system. All elements that support operation create a system of operation (1):

$$SO =< SU, SM, R > \tag{1}$$

Where: SU – usage subsystem (operation); $SU =< \mathbf{UE}, R >$,

$\mathbf{UE} = \{UE_i\}$ – elements of usage subsystem,

SM – maintenance subsystem; $SM =< \mathbf{ME}, R >$,

$\mathbf{ME} = \{ME_i\}$ – elements of maintenance subsystem,

R – relations among system elements.

An object circulates between the operation and maintenance systems so that it requires all necessary resources consisting of an operation and maintenance crew, infrastructure, and environment.

1.2. Operational process of mechanical object

Operation is defined as ”the combination of all technical and administrative actions intended to enable an item to perform a required function, recognising necessary adaptation to changes in external conditions” [3].

An object function that is designed to satisfy a customer’s needs is performed by changes of object states $S = \{s_1, s_2, ..., s_n\}$ in time sequence $t = \{t_1, t_2, ..., t_i, ...\}$. The function describing the distribution of time in states is called the operational process $S(t)$.

The operational process is a subsequent change of the state of the object and, according to the main function of the object, it is recognised as in an *up* or *down* state. The process jumps between up and down states in random moments. Failures and repairs (Fig. 1) determine the instant of jump.

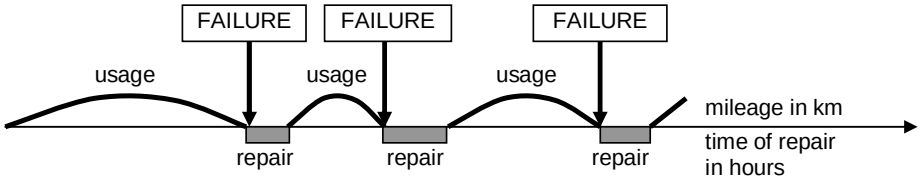


Fig. 1 Operational process of repaired object
Rys. 1. Proces eksploatacji obiektu naprawialnego

In an operational process, one may distinguish controlled and uncontrolled processes. Controlled processes are planned by man/management, depend on required tasks and management methods, and are usually more or less predictable. Timetables, schedules, plans of usage or scheduled maintenance are controlled processes. All unpredictable events that disturb the above processes sometimes make these processes uncontrolled. We are usually forced to weather catastrophes (storms, hurricane winds, heavy snow, flood, etc.), technical catastrophes (crashes, collisions, building, bridges or machines collapses, explosions or fires) or human errors while operating an object.

Another classification criterion due to process definition is availability. An object being in the state in which it cannot perform a desired function is in fault state [3].

The operational process starts by introducing an object in operation (purchase and installation) and finishes by withdrawal from operation.

In a practical approach, there are two kinds of decisive events for the moment of decommissioning of the object from operation:

- Random events of catastrophic/disaster character causing the destruction of the object, and
- Purposeful operational decisions concerning a withdrawal from the use or thorough reconstruction of the object.

A catastrophe is an event during which the destruction of the supporting structure and of the majority of sets and assemblies being essential for fulfilling the functions of the object takes place. Generally, the result of a disaster is a withdrawal of the object from operation.

A decision concerning a thorough reconstruction (modernisation) or withdrawal from use is the consequence of diagnostic investigations and of an economical analysis.

These analyses determine further worthwhile and safety aspects of the object being operated.

Safety is one of the most essential criteria, because events resulting in losses of human life constitute inadmissible object behaviour during operation and are classified by the European Organisation of Quality Control among critical object features.

In this connection, object degradation influences the object history in an uncovered or a hidden manner. An open degradation image is observable by means of all kinds of diagnostic examinations, from simple organoleptic inspection to advanced measurement, metallographic, X-ray, gammascopy techniques, etc. The state of the object (its degradation degree), determined on the basis of evaluation measures according to assumptions, e.g. the total degradation degree q_{Σ}^{T} [4], permits one to define the decrease of its operation potential and residual life. In this case, a data bank containing information about the object becomes the basis for making operation decisions concerning the future of the object [5, 6].

A hidden degradation of the object can take place in situations of insufficient supervision and of uncontrolled, wasteful exploitation of the object. Then, there exist no procedures forcing continuous or periodical object diagnosing, and the continuing deteriorating technical state of the object can lead to a disaster, in a hidden way and without previous symptoms. In that case, the lack of information about the state of the object does not permit one to determine the time of operation interruption or to proceed to the withdrawal of the object from use. Thus, a lack of information concerning object degradation leads in an inevitable way to a catastrophe [4].

2. Concept of failure and fault in a mechanical object

2.1. Failure as undesired event

The technical state of an object is described by set of selected technical parameters like dimensions, displacement, force, moment, stress, power, velocity, pressure, temperature, etc.

These parameters are designed according to the object functions, requirements and environment. They are kept during the operation (usage) in the assumed range of acceptance [7]. Crossing the limit threshold is equivalent to an event called a failure (Fig. 2).

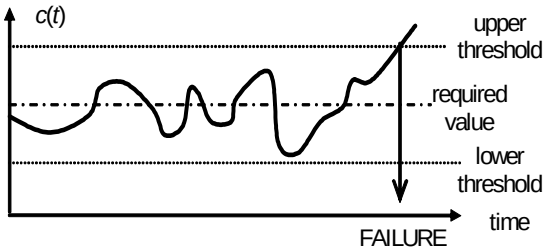


Fig. 2. Variability of technical state parameter and failure moment
Rys. 2. Zmienność parametru stanu technicznego i chwila uszkodzenia

The failure state is defined as the “termination of the ability of an item to perform a required function” [3]. The failure is an event, and after failure the item is in a state called fault.

There are two general approaches to the concept of failure in engineering sciences. The case when parameter $c(t)$ varies randomly according to operational demands or changes monotonically (increases or decreases) (Fig. 3 a and b).

More precise analysis of the failure phenomena shows that the failure as an event occurs when active load exceeds the strength of an object. The safety index represents the ratio of load over strength and assures that, at the design stage, with some level of confidence, the undesired event (failure) should not happen in the real world.

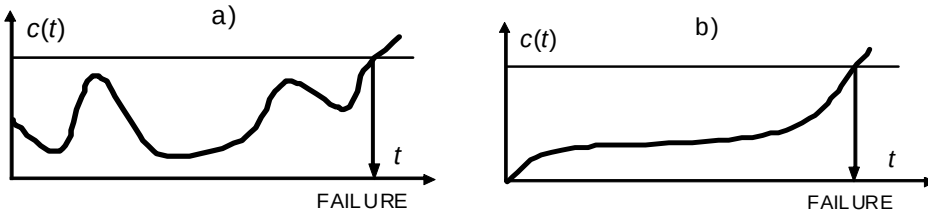


Fig. 3. Behaviour of technical parameter in operation,
a) random variation of parameter $c(t)$, b) monotonic increase of parameter $c(t)$ value
Rys. 3. Zachowanie parametru technicznego w eksploatacji
a) przypadkowa zmienność parametru c(t), b) monotoniczny wzrost wartości parametru c(t)

In fact, in real operation, both load and strength may be regarded as random processes and static reliability is defined [8] as the probability that the current load does not exceed the strength of the element (2):

$$R = P(L < S) = \int_0^{\infty} F_L(s)\, f_S(s) ds \tag{2}$$

Where: R – probability of safe relation between load L and strength S ($L<S$),
F_L – distribution function of load,
f_S – density function of strength.

If load and strength are both normally distributed, respectively, $N(\overline{L}, \sigma_L)$ and $N(\overline{S}, \sigma_S)$ as shown in Fig. 4, than the safety margin SM is calculated as

$SM = \overline{S} - \overline{L}$. Applying $\sigma_m = \sqrt{\sigma_S^2 + \sigma_L^2}$, the reliability of an item is then defined as $P\left(\frac{\overline{S} - \overline{L}}{\sqrt{\sigma_S^2 + \sigma_L^2}}\right) > 0$ [8, 9].

Variability of technical state parameters may take place regarding the internal strength of the object as well as the external load applied during operation. These processes are usually classified as two main types, presented in Fig. 4. It gives four pictures of failure as a combination of the variability of strength and load. The case shown in Fig. 5d corresponds to level crossing with random bound [8].

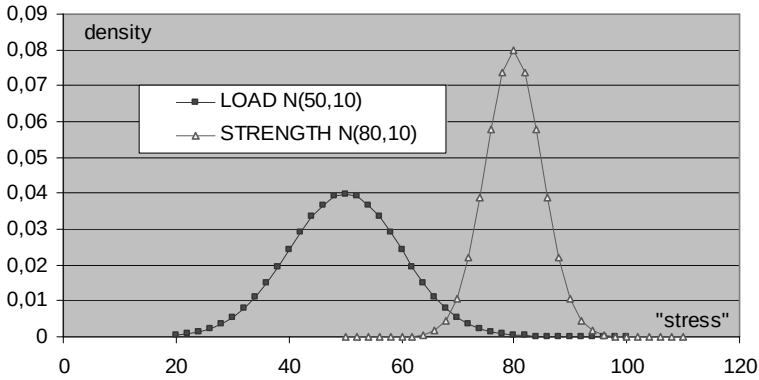


Fig. 4. Example of the relation between load and strength normally distributed
Rys. 4. Przykład relacji między obciążeniem i wytrzymałością opisanymi rozkładem normalnym.

The degradation of a technical object (deteriorating of strength) is a phenomenon consisting in the loss of its usability potential and is described as a stochastic process with respect to the real time of operational use. Degradation depends upon the lapse of time and operational and environmental conditions [10,4]. The object's technical state q can be described as a vector $\mathbf{q}(t)$ of selected criterion parameters $c_i(t)$:

$$\mathbf{q}(t) = \langle c_i(t) \rangle, \quad i = 1, I; \quad c_i(t) \in \chi_i(t) \tag{3}$$

which determines the instantaneous abilities of the object to perform assumed functions [11], [12]. Thus, object availability is a state in which each of the criterion parameters is included within intervals of admissible variability $C(t) = \langle c_i^{\min}(t), c_i^{\max}(t) \rangle$ (Fig. 2). That means, in the traditional damage model, that an excess of admissible values of at least one of the distinguished

parameters is equivalent to the damage of the object and to its passage to the fault state (Fig. 5). With reference to real operational use, the model of unavailability can be generalised through an expansion of the area of technical criterion parameters by economical, safety, environment protection, and other criteria.

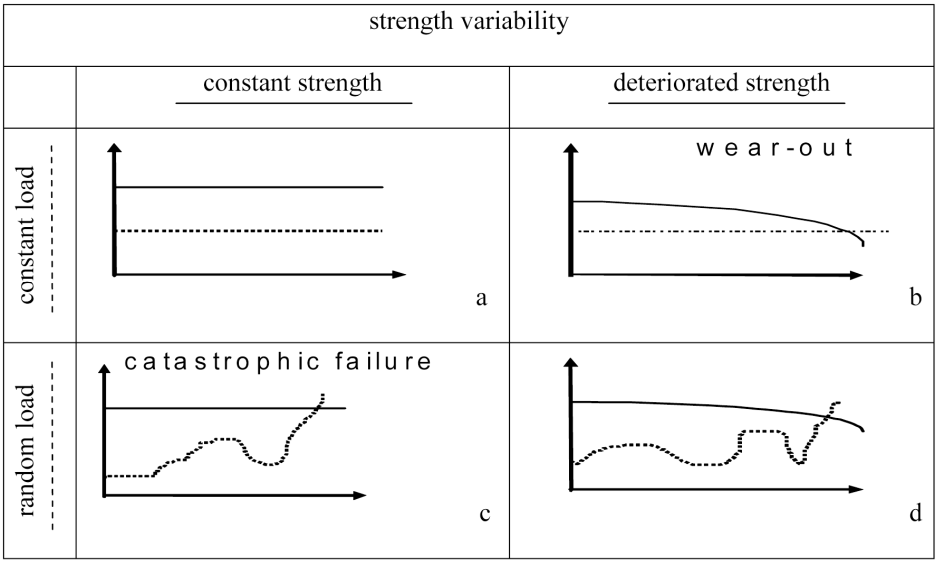


Fig. 5. Combinations of load and strength in operation
Rys. 5. Kombinacje obciążenia i wytrzymałości w eksploatacji

2.2. Typical fault modes

Fault mode is “one of the possible states of a faulty item” [3] and it is how we observe a consequence of a failure. It is the way of demonstrating the inability of performing a function like rapture, bend, fracture, seizing, wear, and many others. Physical processes that lead to failure classify fault models in two groups: “wear out” and “overstress” models (Fig. 3) [12].

The most typical wear out failures in mechanical components are wear, fatigue, creep and corrosion, but there are also observed failures that are a combination of the mechanisms mentioned above like stress and electrochemical corrosion or degradation in strength due to stress variability or high temperature. It is also necessary to mention an influence of man as a failure cause. It is believed that about 80% of failures are introduced by operators or maintenance crews [13]. The wear out observed in a life time creates an increasing/decreasing monotonic process so that variables $c_i(t)$ reach a threshold limit value at some time.

Fig. 6 shows the simplest examples of wear-out and sudden failure. The complete loss of friction in a brake shoe block is probably due to poor maintenance. A connecting rod is torn by tensile impact during piston seizing (to the right).

An example of fatigue failure with a characteristic large fatigue zone, corroded before final fracture, and a glossy, instantaneous fracture zone are shown in Fig. 7.

The most complex failure is represented in Fig. 8. The bearing cap of an engine water pump is broken because of ball bearings released from seized bearing.

Analysis of the fault modes of the entire object usually concerns the inability of the main object to function. Fig. 9 and 10 show a design error resulting in an early crack of the deck transom of a river barge BP-500 [14].

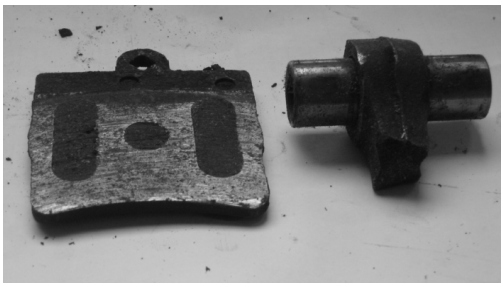


Fig. 6. Example of wear-out failure
(brake shoe- to the left) and sudden disruption
(connecting-rod – to the right)
Rys. 6. Przykład uszkodzenia zużyciowego
(okładzina hamulcowa – po lewej) i nagłego złamania
(korbowód – po prawej)

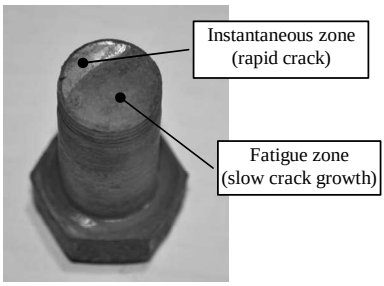


Fig. 7. Fatigue crack: bolt ϕ 24mm.
Rys. 7. Pęknięcie zmęczeniowe śruby
ϕ 24mm

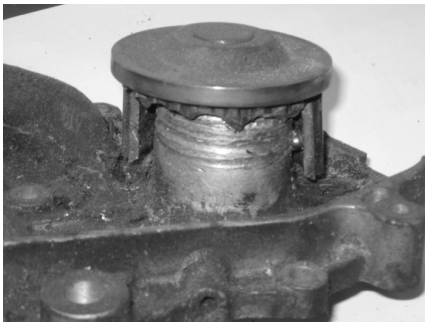


Fig. 8. Total, secondary destruction of water pump (car engine)
due to primary bearing failure
Rys. 8. Całkowite, wtórne uszkodzenie pompy wody
(silnika samochodowego) z powodu pierwotnego uszkodzenia łożyska

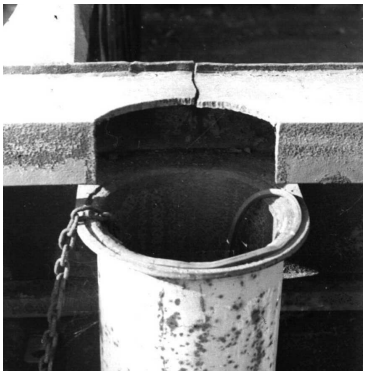


Fig. 9. Crack of deck transom of river barge BP-500 (stress concentration, notch due to design error)
Rys. 9. Pęknięcie pawęży barki BP-500 (koncentracja naprężeń w karbie spowodowanym błędem konstrukcyjnym)

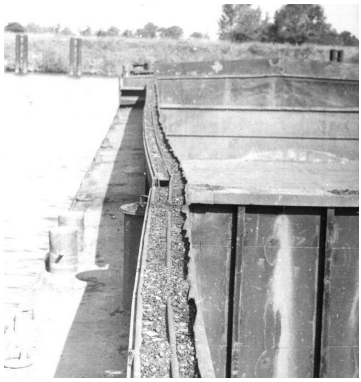


Fig. 10. Example of macro notch of deck transom of river barge BP-500
Rys. 10. Przykład karbu w skali makro pawęży barki BP-500

Pictures above show the variety of fault modes and the necessity of searching for the cause of the failure to prevent future unexpected stops of mechanical objects. Knowledge concerning qualitative and quantitative failure assessment is important in the process of object improvement and modernisation (design) and in setting good operational and maintenance practice.

3. Reliability characteristics of mechanical objects

Reliability in present standards is a part of the wider concept known as *dependability*. It is the collective term used to describe the availability performance and influencing factors: reliability performance, maintainability performance, and maintenance support performance. Dependability is used only for general descriptions in non-quantitative terms [3].

Availability (performance) describes the ability of an item to be in a state to perform a required function under given conditions at a given instant of time or over a given time interval, assuming that the required external resources are provided.

The *reliability* of a product is the probability that the product will perform an expected (designed) function without failure for a given time, at a desired confidence level under specified operating and environmental conditions.

Analysis and assessment of random disturbances of operation process requires working up a reliability model of the object operated in given circumstances. The main factors influencing the variability of failure time are deterioration, ageing, human abilities, and infrastructure conditions [13, 5, 2].

Technical objects, due to failure and repair classification, are described in reliability theory by the following:

- Maintainability, the reparability of an object (model of no repaired or repaired object: repaired with negligible repair time Θ≈0, with any repair time Θ>0),
- Complexity (design, functionality, reliability structure);
- The quantitative assessment of failure (indexes and functions);
- Failure description (cause, mode, consequence, way of repair); and
- Degradation processes analysis setting, for instance, the threshold state of the parameter (ageing, wearing out, fatigue, corrosion, fracturing, etc.).

Randomness of uncontrolled operational processes turns tests, observations and analysis on variables mainly describes the time to or between failures (TTF/TBF) and time to repair (TTR) time for whole objects, its subsystems, assemblies and elements. Statistical process of data concerning TTF and TTR leads to probability distributions and, in consequence, the reliability function R(t), failure distribution F(t), density function f(t), and hazard rate function λ(t).

The classical model of the reliability function is given by Wiener's Formula (4) [10, 11, 12]:

$$R(t) = e^{-\int_0^t \lambda(\tau)d\tau} \tag{4}$$

It combines reliability with hazard rate function $\lambda(t)$, which has a close relation to fault mode (Fig. 11).

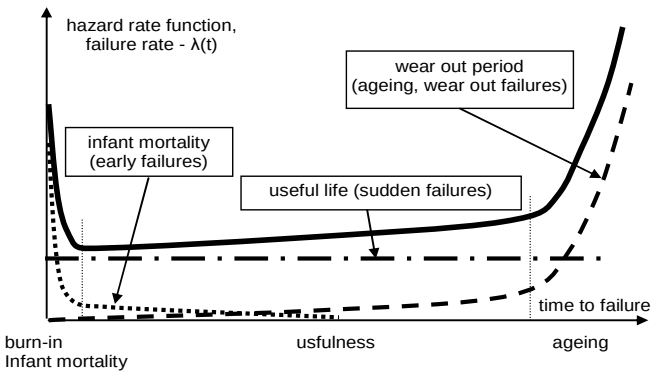


Fig. 11. Hazard rate function and components of bath-tube curve
Rys. 11. Funkcja intensywności uszkodzeń i jej składowe

The relation is bi-directional, i.e., knowing the component fault mode, one may predict the shape of hazard rate function, or, on the other hand, having

calculated theoretical model of the failure, one can show the corresponding failure mode.

An important technical characteristic is B10 (10 percentile), which represents the time to failure (durability), corresponding to 90% certainty that all objects should reach at least time T_{B10} or, in other words, that only 10% of the object may fail before time T_{B10}. Probability distributions taken usually as mathematical models of failure characteristics are Weibull's model, Gauss (normal), log-normal, exponential, beta, and gamma distributions [10,11,5,12].

The examination of failure cause and its mode is of special importance in failure analysis. A high convergence between statistical model of time to failure and failure cause is observed (Fig. 11). Failures caused by natural phenomenon like ageing, wearing or fatigue are described with Gauss distribution (time to failure has normal distribution) with high credibility. Sudden or catastrophic failures caused by reasons external to the object are modelled by exponential distribution [5,12].

4. Application of failure mode knowledge in operation and management

4.1. Automotive spare part stock management

Knowledge of the cause of failure lets us roughly assess the variability of entry to the service stream (service demand). The problem appears in warehouse management, when there are two antagonistic demands. It is necessary to keep in stock large amounts of spare parts to continuously maintain the service process and, on the other hand, cause high reserve expenses for a warehouse. A component of natural or ageing failures (TTF is described by normal distribution) are usually characterised by small variability $\nu = \frac{\sigma_F}{\overline{T}_F} < 0{,}1$, where $\overline{T}_F$ is the mean time between failures and σ_F is the standard deviation of this variable. Components of sudden failures are usually described by exponential distribution and are characterised by large variability $\nu = \frac{\sigma_F}{\overline{T}_F} = 1$, which means that demands on particular components may be expected very rarely as well as very often.

Figs. 12 and 13 show a comparison of distribution functions having the same mean value $\overline{T}_F = 100000$. However, diversification in the standard deviation of normal and exponential distributions makes a great difference in B_{10} index, so that efficient stock for parts with exponential distribution should be much larger. One may observe that, for exponential distribution, 10% of objects will survive a time below 20000 and for normal distribution about 80000

(B10 takes value 80 000 units of time for ageing failures and below 20 000 units for sudden failures). The conclusion is the prediction that spare stock for elements of sudden failures is less anticipated and, to maintain continuity of the maintenance process, should be kept at a higher level.

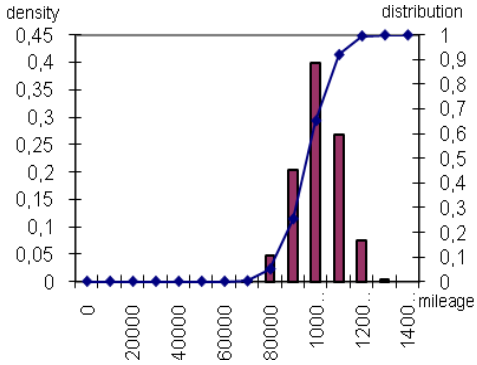


Fig. 12. Density and distribution function as normal distribution with $\overline{T}_F = 100000$ and $\sigma_F = 10000$

Rys. 12. Funkcja gęstości i dystrybuanta rozkładu normalnego $\overline{T}_F = 100000$ i $\sigma_F = 10000$

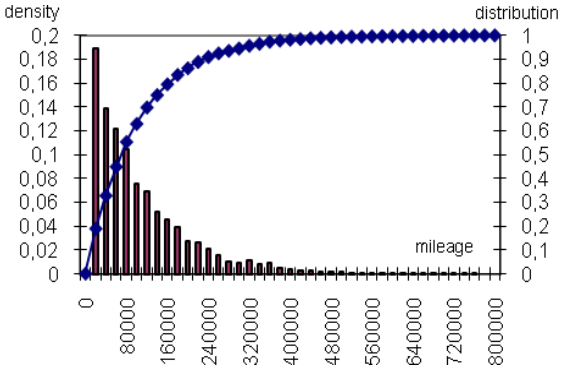


Fig. 13. Density and distribution function as exponential distribution with $\overline{T}_F = 100000$

Rys. 13. Funkcja gęstości i dystrybuanta rozkładu wykładniczego $\overline{T}_F = 100000$ i $\sigma_F = 10000$

The above issue deals only with uniform objects treated individually. In the case of complex objects like vehicle, assemblies, subassemblies of various vehicles, a stock does not undergo the above statement, because it may be mixture of different variables. In that case some asymptotic models are applied.

4.2 Analysis of tank ageing data

Knowledge about reliability characteristics of weapon systems is extremely important in peacetime and during the war [15]. In the period of peacetime, all weapon system are stored or used as training objects. Both in real war service and during peacetime, there is an expected high availability, since they have to provide soldiers safety and fulfil military requirements. Tanks, as main land weapon, should therefore achieve its standard availability as soon as possible while used as training objects.

Reliability tests have been performed on a sample of 144 tanks in a period of over 3 years. The tanks were new, introduced to training system with the manufacturer warranty.

Collected data on 11 functional subassemblies of the tank TWARDY made the evaluation of 11 reliability functions possible. In 6 cases out of 11, a Weibull failure distribution function was obtained with shape parameter scientifically less then 1. It testifies that the period of observation was the burn-in period with the failures the manufacturer's responsibility. Fig. 14 shows the decreasing hazard rate function of the fire control system. In the case of power transmission subassembly, the hazard rate was nearly constant (Fig. 15). It is suspected that failures observed due to that subsystem have the nature of incidents of overloading or human errors while operated by a trainee.

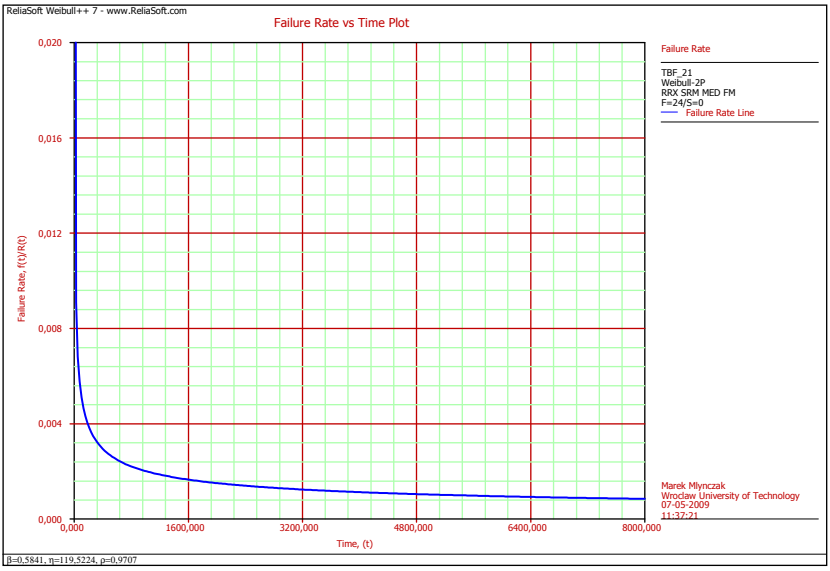


Fig. 14. Function of failure rate (decreasing) of fire control system in tank TWARDY
Rys. 14. Rosnąca funkcja intensywności uszkodzeń systemu uzbrojenia czołgu TWARDY

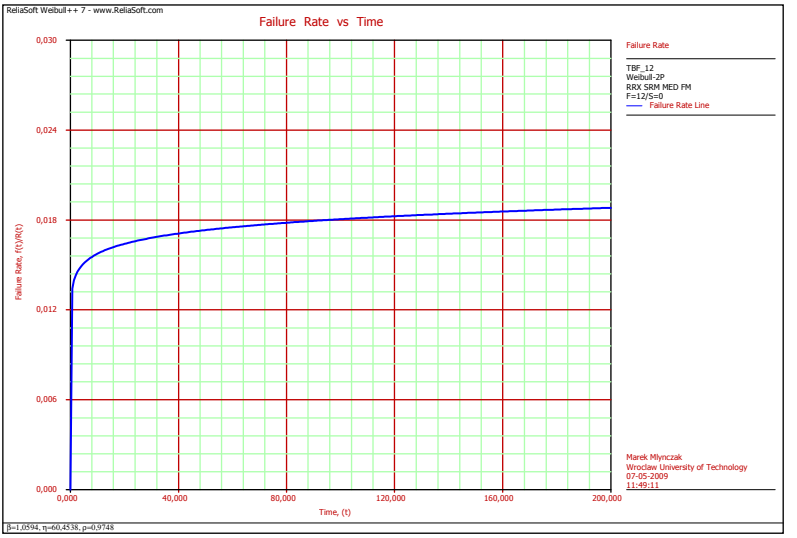


Fig. 15. Function of failure rate (nearly constant) of power transmission system in tank TWARDY
Rys. 15. Bliska stałego przebiegu funkcja intensywności uszkodzeń systemu napędowego czołgu TWARDY

Conclusions

The real operation of mechanical objects provides everyday many examples of failures due to design, manufacture, and operation. Some failures are embedded in the object (hopefully not intentionally), and they appear usually in the beginning of operation process (burn-in failures with decreasing hazard rate function). Long lasting correctly managed operational processes may bring failures of a sudden, catastrophic character related to exponential distribution of time between failures. They are hardly predictable, but the intensities of such events are very low. The last part of an object's life, assuming that it survives to that time, is related to ageing and wear failures due to the degradation of materials of the object. Depredating processes become more rapid with operational time and finally lead to failure. The corresponding failure rate is modelled by a monotonically increasing function. An appropriate mathematical model is Weibull's distribution with a shape parameter larger than 1.0, practically, about 3.3.

References

[1] Augustynowicz J., Dudek D., Dudek K., Figiel A., Młyńczak M, Nowakowski T., Przystupa F.W. (2001) Struktury patologiczne maszyn. KONBiN'2001. Seria Monograficzna ITWL, Szczyrk.

[2] Bentley J.P. (1999). Introduction to Reliability and Quality Engineering. Addison-Wesley Longman Ltd., Edinburgh Gate, Harlow.
[3] Blischke W., Murthy D.N.P. (2000). Reliability. Modeling, Prediction, and Optimization. John Wiley & Sons, Inc. New York.
[4] DUDEK D. (1996). Degradacja maszyn roboczych. Teoria czy sztuka. Problemy Maszyn Roboczych. Z. 7/96. Instytut Technologii Eksploatacji, Radom.
[5] Encyclopædia Britannica® Online. http://www.britannica.com.
[6] Fragola J.R. (2001). Human reliability analysis procedure. Tutorial Notes. Proceedings of the European Conference on Safety and Reliability ESREL'01 Safety & Reliability. Torino.
[7] Gercbach L.B., Kordoński Ch.B. (1968). Modele niezawodnościowe obiektów technicznych. WNT, Warszawa.
[8] Handbook of Reliability Engineering. Ed. Ushakov I.A., Harrison R.A. (1994). John Wiley&Sons Inc. New York.
[9] Hubka V., Eder W.E. (1988). Theory of Technical Systems, A Total Concept Theory for Engineering Design. Springer Verlag. Berlin, Heidelberg.
[10] IEC 60050–191. Dependability and quality of service, Amendments 1 & 2. International Electrotechnical Vocabulary (IEV 191).
[11] Kowalski K., Młyńczak M. (2009). Issue of availability of weapon systems in early operation phase. MOTROL, Lublin.
[12] Młyńczak M. (1999). Maintenance modeling of degrading objects. Proceedings of ESREL'99. A.A. Balkema Monachium.
[13] O'Connor P.D.T. (1985). Practical Reliability Engineering. John Wiley & Sons. A Wiley-Interscience Publication. Chichester, New York, Brisbane, Toronto, Singapore.
[14] Smalko Z. (1972). Podstawy projektowania niezawodnych maszyn i urządzeń mechanicznych. WNT, Warszawa.
[15] Smith D.J. (2007). Reliability, Maintainability and Risk. Oxford: Elsevier.

Manuscript received by Editorial Board, February 11th, 2011

Modele uszkodzeń obiektów mechanicznych

S t r e s z c z e n i e

Obiekty mechaniczne są eksploatowane w rzeczywistych warunkach, gdzie głównymi czynnikami wpływającymi na ich efektywność są zjawiska degradacji materiału elementów obiektu oraz zmienność procesu eksploatacji sterowanego przez operatora. Degradacja jest długotrwałym procesem dotyczącym materiału struktury konstrukcyjnej oraz połączeń elementów. Utrata zdatności użytkowej maszyny wynika z uszkodzeń spowodowanych tarciem i zużyciem, zmęczeniem, korozją, przeciążeniem, starzeniem materiału i innymi procesami destrukcyjnymi. Można zauważyć bliski związek między postaciami uszkodzeń a modelami niezawodności. Wiedza o uszkodzeniach wspomaga analityków w tworzeniu modeli niezawodnościowych i wskazuje optymalne decyzje eksploatacyjne. W pracy omówiono związki między zjawiskami fizycznymi i teoretycznymi modelami tworzące wspólną platformę dla procesu decyzyjnego.

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

TOMASZ NOWAKOWSKI*,
SYLWIA WERBIŃSKA-WOJCIECHOWSKA**

Maintenance models – role of logistic support review

Key words

Logistic support system, maintenance system modelling, spare part inventory models.

Słowa kluczowe

System wsparcia logistycznego, modelowanie systemów obsługi, modele zapasów części wymiennych.

Summary

This paper presents an overview of some recent developments in the area of inventory planning and maintenance scheduling issues. The emphasis is on spare part inventory models, which the authors divided into four main groups of models: models of optimal spare part inventory policy for system under PM, number of spare parts optimisation models, storage reliability models, multi-echelon systems models. The first group of models is discussed in more depth. Later, the paper considers the time dependent system of systems where the system total task must be executed during the constrained time resource. The simulation approach is used to investigate the influence of parameters of the procurement process (order quantity, critical inventory level, lead-time length) on the system of systems behaviour.

* Wroclaw University of Technology, Faculty of Mechanical Engineering, Institute of Machine Designing and Operation, 50-371 Wrocław, ul. Łukasiewicza 5; phone: +48 71 320-35-11, e-mail: tomasz.nowakowski@pwr.wroc.pl

** Wroclaw University of Technology, Faculty of Mechanical Engineering, Institute of Machine Designing and Operation 50-371 Wrocław, ul. Łukasiewicza 5; phone: +48 71 320-37-67, e-mail: sylwia.werbinska@pwr.wroc.pl

1. Introduction

The interest in the development and investigation of maintenance problems has been extensively discussed in literature since the early 1960s. The basic review in the area of maintenance modelling is prepared by Pierskalla & Voelker [1], where the authors investigated discrete time vs. continuous time maintenance models, later updated by Valdez-Flores & Feldman [2]. For other surveys see e.g. [3–11].

However, most of the maintenance models investigated in the literature on reliability theory assume that all the necessary logistic support resources, which include maintenance resources, support personnel, logistic information and data, spare and repair parts, and facilities, are immediately provided when desired. In practice, the repair capacity is not infinite, and logistic information may be unreliable. Moreover, the influence of a spare provisioning policy on the maintenance policy also cannot be ignored, since spare parts are ordered and carried in a limited quantity and the procurement lead-time is not negligible [11].

Throughout the years, the importance of the logistic support functions and logistic support management has grown. The plethora of studies, which have addressed the problem of logistic support systems modelling, can be divided into four main groups being presented in Fig. 1. A bibliography of the work done can be found in [11].

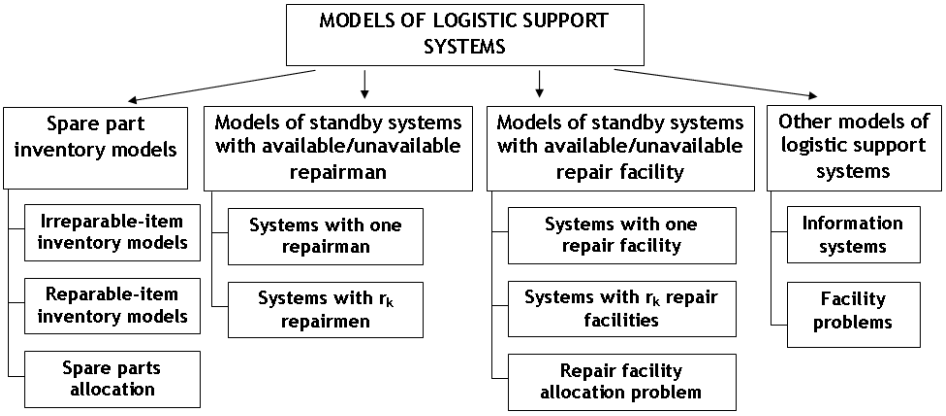


Fig. 1. Classification scheme of logistic support system models [11]
Rys. 1. Klasyfikacja modeli wsparcia logistycznego funkcjonowania systemu technicznego [11]

The problem of providing an adequate and efficient supply of spare parts, in support of maintenance and repair of operational systems, has been researched for many decades.

A sufficient inventory level of spare parts has to be maintained to keep the system in operating condition. When the spare elements are under stocked, it may lead to extended system downtime and, as a consequence, have a negative impact on the system availability level. On the other hand, maintenance of excessive spare parts can lead to large inventory holding costs. Moreover, the requirements for planning the logistics of spare elements necessary in maintenance actions of operational systems differ from those of other materials in several ways: the service requirements are higher, the effects of stock-outs may be financially remarkable, the demand for parts may be sporadic and difficult to forecast, and the prices of individual parts may be very high.

Consequently, one of the most important decisions faced by maintenance managers is the determination of optimal stocking levels, which involves finding the answer to the following questions, such that the total expected inventory costs be minimised:

- When to (re)order?
- How many items to (re)order?

There are many models in the literature on reliability theory regarding spare parts supply process optimisation. A significant portion of them are bases on a classic inventory theory, where the procurement process parameters are usually optimised taking into account the cost constraints (see e.g. [12]).

Many inventory papers that treat stock replenishment problems for stochastically failing equipment/systems are surveyed in [1] and updated by [3, 13, 14].

Following the introduction, this paper is focused on spare part inventory optimisation problems. The paper is organised as follows: In the Section 2, we present an overview of the most often applied models. We focus on one group of defined models. For more detailed investigation of other groups of models see [6, 11]. Later, we provided an example of a time dependent system of systems, where the system total task must be executed during the constrained time resource and briefly summarise.

We used a “system of systems” concept to model the interactions between the operational system and its logistic support system. According to the definition [15], *the system of systems context arises when a need or a set of needs are met with a mix of multiple systems, each of which are capable of independent operation but must interact with each other in order to provide a given capability*. More information can be found in [16, 17, 18].

2. Spare part inventory models

The reparable-item inventory problem has received much attention in the logistics literature. The general classification scheme for spare part inventory models is presented in Fig. 2.

The models presented below regard single-echelon systems. However, for advanced technical systems, such as engines, or aeroplanes, high system availability is enhanced thanks to the multi-echelon inventory system, in which usually two or more echelons are equipped with repair and stocking facilities. For more information see e.g. [11, 19, 20]

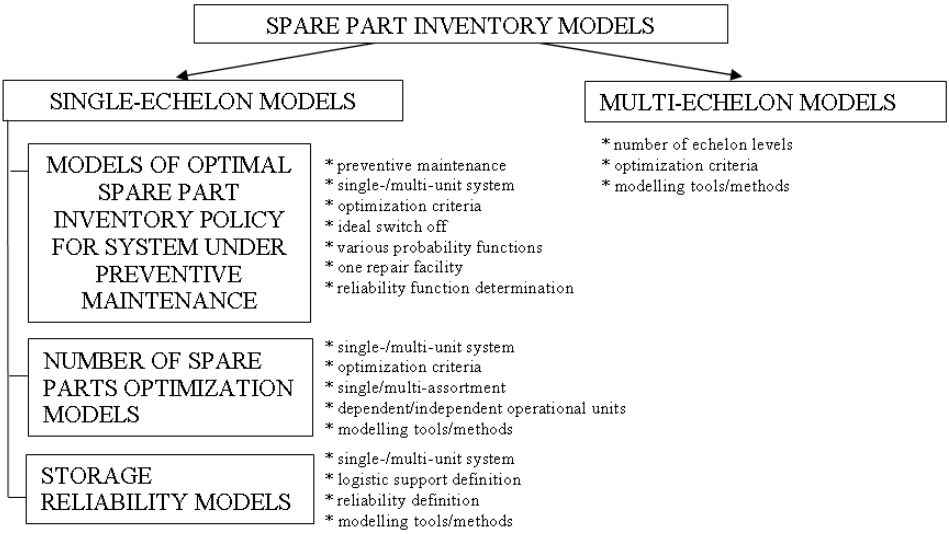


Fig. 2. Spare part inventory models classification [11]
Rys. 2. Klasyfikacja modeli sterowania zapasami obiektu technicznego podlegającego procesowi obsługiwania [11]

2.1. Models of optimal spare part inventory policy for a system under preventive maintenance

The main classification scheme of the investigated models is presented in Fig. 3.

There are two fundamental types of maintenance – preventive maintenance (PM) and corrective maintenance (CM). For PM demand for spare parts is predictable. For such maintenance, it may be possible to order parts to arrive just in time for use, and it may not be necessary to stockpile repair parts at all. In case of unplanned repair, the consequences of stock-outs regard to system unavailability has significant costs, and that is why some kinds of stock policies are necessary. It is natural in technical systems that only spare units that can be delivered by order are available for maintenance/replacement. In this case, we cannot neglect a lead-time for delivering the spare unit.

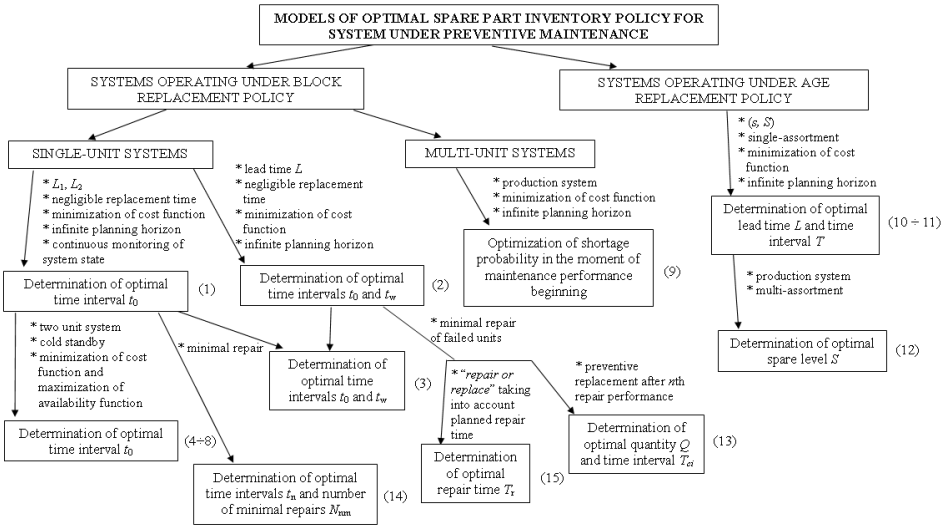


Fig. 3. Models of optimal spare part inventory policy for system under PM [11]
Rys. 3. Modele doboru optymalnych parametrów strategii sterowania zapasami, gdy strategia obsługi profilaktycznej obiektu jest zdefiniowana [11]

The first models investigated the possibility of spare part shortage due to delivery process performance regarded single-unit systems (see e.g. [21, 22]). In [21], the authors consider a one-unit system, where each failure is scrapped without repair and each spare part is only provided after a lead-time by an order. In the presented model, the following policy is adopted: An order for a spare is made at a pre-specified time instant t_o during an operating period of an original unit that is called a regular order. The lead-time entails L_2 time units. After delivery, the original unit is replaced whether it is operable or not. However, if the failure of the unit takes place before the time instant t_o, an emergency order is made at a failure time immediately. After an emergency delivery, which entails L_1 time units, a failed unit is replaced, and the process repeats itself.

Taking into account the following assumptions:

- Infinite planning horizon;
- Negligible replacement time of operating unit; and,
- The system is continuously kept under constant observation until a pre-specified time instant t_o or till the instant of failure, whichever occurs first.

The expected cost per unit time in the steady state is given by the following formula:

$$C_s(t_o) = \frac{c_m \int_0^{t_0} R(t)dt + c_{p1}F(t_0) + c_{p2}R(t_0) + c_{dw}\left[(L_1 - L_2)F(t_0) + \int_{t_0}^{t_0+L_2} F(t)dt\right]}{(L_1 - L_2)F(t_0) + L_2 + \int_0^{t_0} R(t)dt} \quad (1)$$

Where: c_{p1} – cost of spare element expedited order which is made at time instant t,
c_{p2} – costs of spare element regular order made at time t_o,
c_{dw} – cost of system downtime per unit time,
$L_1(L_2)$ – random lead time for emergency (regular) order,
c_m – cost of system observation proportional to the expected duration of observation,
$R(t)$ – system reliability function,
$F(t)$ – cumulative distribution function of unit.

The presented model development can be found in [22], where the additional assumption is made that the operational unit replacement is made in one of two situations, whichever occurs first: when unit fails or when time of PM occurs at time instant t_w. The expected cost per unit time in a steady state and is given by the following:

$$C_z(t_0,t_w)=\frac{c_{p1}F(t_0)+c_{p2}R(t_0)+c_{dw}\int_{t_0}^{t_0+L}F(t)dt+c_h\int_{t_0+L}^{t_w}R(t)dt}{\int_0^{t_w}R(t)dt+\int_{t_0}^{t_0+L}F(t)dt} \tag{2}$$

Where: c_h – cost of holding a spare unit in a stock per unit of time,
L – random lead time.

Another work, made by Dohi et al. [23], presents a generalised order-replacement model arising in spare part inventory management, which is based on the assumptions taken in [21] and [22]. A replacement problem is considered for one-unit systems where each failed unit is scrapped and each spare part is provide, after a lead-time, by an order. If the original unit does not fail up to a pre-specified time instant t_o, the regular order is made, and after a lead-time L_2 the spare unit is delivered. The delivered spare element is put into inventory until the moment of the original unit failure or until the moment of PM, whichever occurs first. On the other hand, if the original unit fails before the time instant t_o, the expedited order is made immediately at the failure time, and the spare part takes over its operation just after it is delivered after a lead time L_1. In this situation, the regular order is not made. The function of the expected total cost per time unit is given by the following:

$$C_z(t_0,t_w)=\left\{c_{sp}\left\{\int_0^{\infty}\int_0^{t_0}L_1dF(t)dE_1(L_1)+\int_0^{\infty}\int_{t_0}^{t_0+L_2}(t_0+L_2-t)dF(t)dE_2(L_2)\right\}+c_{p1}F(t_0)+\right.$$
$$+c_{p2}R(t_0)+c_h\int_0^{\infty}\int_{t_0+L_2}^{t_0+L_2+t_w}(t-t_0-L_2)dF(t)dE_2(L_2)+$$
$$\left.+c_h\int_0^{\infty}\int_{t_0+L_2+t_w}^{\infty}t_wdF(t)dE_2(L_2)\right\}\times\left(\int_0^{\infty}\int_0^{t_0}(t+L_1)dF(t)dE_1(L_1)+\int_0^{\infty}\int_{t_0}^{t_0+L_2}(t_0+L_2)dF(t)dE_2(L_2)+\right. \tag{3}$$
$$\left.+\int_0^{\infty}\int_{t_0+L_2}^{t_0+L_2+t_w}tdF(t)dE_2(L_2)+\int_0^{\infty}\int_{t_0+L_2+t_w}^{\infty}(t_0+L_2+t_w)dF(t)dE_2(L_2)\right)^{-1}$$

Where: c_{sp} – shortage cost per unit of time,
$E_i(t)$ – cumulative distribution function of L_i (i = 1, 2).

The problem of optimal spare part ordering policies for two-unit cold standby redundant system with two dissimilar units is considered in [24]. The replacement problem is defined as follows: Unit 1 begins working and unit 2 is in standby at time 0, and the planning horizon is infinitive. If unit 1 does not fail up to a pre-specified time t_o, the regular order for spare parts of both units 1 and 2 is made at time t_o. After a lead-time L_2 the spare parts are delivered, and at the time t_o+L_2, all original units are replaced correctively/preventively by spares, irrespective of the states of original ones, since the two units are not identical. The order for two spare parts is always needed. On the other hand, if the unit 1 fails before the time t_o, the operation is switched to the unit 2, and the expedited order for spare parts of both units is immediately made at the failure time. All original units are replaced by spares just after delivery, which lasts a lead-time L_1. The switchover is assumed to be perfect and instantaneous. The state in which both units fail before delivery of spare units implies the system down.

To obtain the optimal ordering policy parameter, the expected cost per unit time in the steady state and the stationary availability are developed.

The expected inventory cost function for one cycle is given by the following:

$$C_z(t_0)=c_{sp}\left\{\int_0^{t_0}\int_0^{L_1}(L_1-s)dF_2(s)dF_1(t)+\int_{t_0}^{t_0+L_2}\int_0^{t_0+L_2-t}[(t_0+L_2)-(t+s)]dF_2(s)dF_1(t)\right\}+$$
$$+c_{rz}\left\{\int_0^{t_0}\int_{L_1}^{\infty}(s-L_1)dF_2(s)dF_1(t)+\right. \tag{4}$$
$$\left.+\int_{t_0}^{t_0+L_2}\int_{t_0+L_2-t}^{\infty}[(t+s)-(t_0+L_2)]dF_2(s)dF_1(t)+\int_{t_0+L_2}^{\infty}[t-(t_0+L_2)]dF_1(t)+\int_{t_0+L_2}^{\infty}\int_0^{\infty}sdF_2(s)dF_1(t)\right\}+$$
$$+c_h\left\{\int_0^{t_0}tdF_1(t)+\int_{t_0}^{t_0+L_2}tdF_1(t)+\int_{t_0+L_2}^{\infty}(t_0+L_2)dF_1(t)\right\}+c_{p1}F_1(t_0)+c_{p2}R_1(t_0)$$

Where: c_{rz} – cost per unit time incurred for the residual lifetime of the original unit, which is still operable.

The expected time for one cycle is defined as:

$$E[T_{c_j}(t_o)] = \int_0^{t_0} (t + L_1) dF_1(t) + \int_{t_0}^{\infty} (t_0 + L_2) dF_1(t) \tag{5}$$

Applying the renewal theorem, the expected cost per unit time in the steady state is given by the following:

$$C_s(t_o) = \frac{C_{sj}(t_o)}{E[T_{cj}(t_o)]} \tag{6}$$

Moreover, the stationary availability $A(t_o)$, defined as the probability that a system is operative in the steady state, is given by the following:

$$A(t_0) = \frac{E[T_{oj}(t_0)]}{E[T_{c_j}(t_0)]} \tag{7}$$

Where: $T_{oj}(t_o)$ – effective time of a system for j-th cycle, given by the following formula.

$$\begin{aligned} E[T_{o_j}(t_0)] = & \int_0^{t_0}\int_0^{L_1} (t+s) dF_2(s) dF_1(t) + \int_0^{t_0}\int_{L_1}^{\infty} (t+L_1) dF_2(s) dF_1(t) + \int_{t_0}^{t_0+L_2}\int_0^{t_0+L_2-t} (t+s) dF_2(s) dF_1(t) + t_0 + \\ & + \int_{t_0}^{t_0+L_2}\int_{t_0+L_2-t}^{\infty} (t_0+L_2) dF_2(s) dF_1(t) + \int_{t_0+L_2}^{\infty} (t_0+L_2) dF_1(t) \end{aligned} \tag{8}$$

Another very interesting problem regarding spare parts inventory planning in order to keep a production system in operating condition is considered. An example of such a system, consisting of n identical and stochastically independent production machines in k-out-of-n reliability structure, is given in [25]. The operational process of the presented system includes planned machine shutdown, during which all of the failed elements in that maintenance cycle are replaced in order to increase system reliability.

The problem considered in the presented paper regards the a priori planning of spare part inventories required for maintenance during a phased mission. In the system, it is possible to replace failed elements only during overhauls performed between two phases. The replacement of the failed elements at the end of phase k may be done by spare parts remaining unused from the proceeding overhauls and by S_k items planned to become available at time point t_k. the failed elements, after replacement, are repaired and put into stock (inventory with returns system). The shortage can occur when the demand exceeds the number of spare elements available from stock, and then spare parts

are obtained by an emergency order or by borrowing, a penalty cost is paid, and the mission continues. The problem of spare part planning is to find S_k for which stock out of probabilities p_{spk} at time point t_k are smaller than the specified numbers α_{spk}:

$$\begin{cases} \min C_z = \sum_{k=1}^{K} c_z S_k \\ p_{spk}(S_1,...,S_K) \le \alpha_{spk} \end{cases} \tag{9}$$

Where: C_z – function of expected total purchase and holding cost,
c_z – total purchase and holding cost per unit per unit of time,
t_k – moments of planned overhauls, $k = 1, 2, .., K$,
S_k – planned number of spare elements at t_k,
$p_{spk}(S_1,..., S_K)$ – probability of stock out at t_k,
α_{spk} – maximal level of stock out probability in one maintenance cycle.

The model is solved with the use of a Markov process whose states are determined by the number of available spare parts and the following assumptions:

- Perfect maintenance conditions, and
- Elements of the system are identical and identically distributed.

Many works that address the problem of determining the optimal ordering policy parameters for technical systems operating under a block replacement policy are based on using simulation processes (see e.g. [26, 27, 28]). Those models give the solutions to define optimal ordering policy parameters (e.g. order placement moments) and define optimal PM parameters (moments of maintenance performance).

Another group of models where the problem of spare inventory optimisation is investigated regards the age replacement policy. The problem of age replacement policy with inventory restrictions can be found in, e.g. [29], where authors investigated two inventory policies (s,S).

According to the (1,1) inventory policy, operating element will be replaced in one of two situations: at age T or whenever the minimal repair cost C_{nm} is greater than some predetermined value $C_{nm}^{\max}$, whichever occurs first. When the unit must be replaced, it will be ordered and delivered after a lead-time L. During the time of waiting for spare part delivery, the system is in downstate.

Optimisation of the following parameters: lead time L and the age T when a system must be replaced base on the minimisation of total maintenance and inventory cost, is defined by the following formula:

$$C_s(L,T)=\frac{c_d(L)+\left(c_{w1}-c_{w2}+\int_0^{C_{nm}^{br}} x dF_{cn}(x)\left(1-F_{cn}(C_{nm}^{\max})\right)^{-1}\right)\left[1-\left(R(t)\right)^{1-p_{cnm}}\right]+c_{w2}}{L+\int_0^T \left(R(u)\right)^{1-p_{cnm}} du} \tag{10}$$

Where: $F_{cw}(x)$ – cost of repair distribution function,

p_{cnm} – probability that defined minimal repair cost is greater than $C_{nm}^{\max}$,

$c_d(L)$ – function of costs associated with delivery performance (e.g. ordering cost, cost of lost production).

This kind of model might correspond to some very critical but expensive piece of equipment where one backup is provided [29].

The second investigated inventory policy called (2,2) is a modification of described (1,1) inventory models. In this model, the system will always contain one unit in operation and one additional unit either in inventory or on order. According to the model assumptions, when an operating unit fails, one of two possible situations can happen:

- An additional unit being in inventory is immediately available for replacement,
- An additional unit is on delivery – then the failed unit is repaired at all cost.

No system downtime is ever allowed.

The total cost function is defined as follows:

$$C_s(L,T)=\left\{c_d(L)-\ln R(L)C_{nm}^{L}+\left[1-\left(R(L)\right)^{1-p_{cnm}}\left(R(T)\right)^{1-p_{cnm}}\right]\times\left(\int_0^{C_{nm}^{br}} x dF_{cn}(x)\left(1-F_{cn}(C_{nm}^{\max})\right)^{-1}+c_{w1}-c_{w2}\right)+c_{w2}\right\}\times$$
$$\times\left(L+\left(R(L)\right)^{1-p_{cnm}}\int_L^T \left(R(\tau)\right)^{1-p_{cnm}} d\tau\right)^{-1} \tag{11}$$

Where: C_{nm}^{L} – expected repair cost function during a lead time.

For more complicated problem investigation (see e.g. [30]) simulation processes, dynamic programming, integer programming, and nonlinear programming are the main tools suggested.

Lots of models for the joint optimisation of an optimal age-dependent inventory policy and PM policy regarding production systems subjected to random machine breakdowns (see e.g. [31, 32]) have been investigated. An interesting inventory problem is investigated in [33], where authors developed the optimal number of inventories $S = S_1, S_2, \ldots, S_r$ when

- The system performs under age replacement policy, and
- The system failure rate increase with its age.

The optimisation problem is stated as follows:

$$\begin{cases} \max \prod_{q=1}^{n_q} p_q(S_q) \\ \sum_{q=1}^{n_q} H_q s_q \le C_z^{\max} \end{cases} \tag{12}$$

Where: n_q – number of types of spare elements in production machine,
$p_q(S_q)$ – probability that there will be no stock out of spare elements type q during the overhaul,
s_q – initial inventory level,
$C_z^{\max}$ – maximal allowed level of inventory costs.

The solution of the stated optimisation problem is received with the use of dynamic programming.

However, in real life systems, the failed element can be replaced or repaired, which needs to give an answer to the following questions:

- When should the unit be repaired instead of replaced?
- How many spare parts should be ordered in order to meet demand while keeping the ordering and inventory costs minimal?

One of the models that tries to answer these questions is presented in [34]. In this paper, a joint stocking and replacement model with minimal repair at failure is considered. The authors assumed that Q units are purchased per order, operation unit is replaced after using for time interval T_{ci}, if inventory level is (i-1), and minimal repair is performed for any intervening failures. The problem is to select optimal order quantity Q and replacement intervals T_{ci}, so as to minimise the total maintenance and inventory cost per unit time, given by the following formula:

$$C_s(Q,T_{ci}) = \frac{c_o + c_{wz}Q + c_{nm}\sum_{i=1}^{Q} H(T_{ci}) + c_h \sum_{i=1}^{Q}(i-1)T_{ci}}{\sum_{i=1}^{Q} T_{ci}} \tag{13}$$

Where: c_o – cost of an order placing,
c_{wz} – replacement cost per unit,
i – inventory level.

The investigated problem of optimal ordering and maintenance policy parameter definition is continued in many recent papers (see e.g. [35]). In this paper, the authors defined optimal ordering point t_o and the optimal number of minimal repair N_{nm} before PM performance for a single-unit system. Assumptions made in this model are the same as presented in [21].

An order for a spare element is placed before the nth failure of the operational unit occurrence (moment t_n), and lasts a lead time L. If the operational unit fails before t_n moment occurrence, the system is in a downstate until the moment t_n+L. However, if the unit fails after spare element delivery (kth failure), the system is immediately replaced. Other failures are minimally repaired in time $(0, t_k)$.

Optimisation of parameters is performed with the minimisation of the total expected cost function given by the following:

$$\begin{aligned} C_s(t_n, N_{nm}) = \Bigg\{ & (N_{nm}-1)c_{nm} + c_o + \left(\int_0^\infty \int_x^{x+L} f_{t_n,t_k}(x,y)dydx \right) \times \left(c_{sp} \int_0^\infty \int_x^{x+L} (x+L-y) f_{t_n,t_k}(x,y)dydx \right) + \\ & + \left(\int_0^\infty \int_{x+L}^\infty f_{t_n,t_k}(x,y)dydx \right) \left(c_h \int_0^\infty \int_{x+L}^\infty (y-x-L) f_{t_n,t_k}(x,y)dydx \right) \Bigg\} \times \\ & \times \Bigg\{ \left(\int_0^\infty \int_x^{x+L} f_{t_n,t_k}(x,y)dydx \right) \left(\int_0^\infty \sum_{j=1}^{t_n-1} \frac{\exp(-R(x))(R(x))^j}{j!} dx + L \right) + \\ & + \left(\int_0^\infty \int_{x+L}^\infty f_{t_n,t_k}(x,y)dydx \right) \left(\int_0^\infty \sum_{j=1}^{k_{nm}-1} \frac{\exp(-R(x))(R(x))^j}{j!} dx \right) \Bigg\}^{-1} \end{aligned} \tag{14}$$

Where: $f_{t_n,t_k}(x,y)$ – probability density function of t_n and t_k.

Another interesting solution of the problem ‘replace or repair’ is given in [36], where a simple repair-time limit replacement problem with imperfect repair is considered. The authors investigated a single-unit system in which, when the unit fails, one estimates the repair time. If the repair can be completed up to a pre-specified time limit $T_{r\max}$, the repair is started immediately, otherwise, the spare unit is ordered with a lead-time L. The expected total cost per unit time in the steady state is given by the following formula:

$$C_s(T_r) = \frac{(k_f + c_n)\int_0^{T_r} t dG(t) + (k_f L + c_o)\overline{G}(T_r)}{\int_0^{T_r} t dG(t) + \frac{1}{\lambda} + \left(L + \frac{1}{\lambda^s} - \frac{1}{\lambda} \right)\overline{G}(T_r)} \tag{15}$$

Where: k_f – penalty cost per unit time when system is in downstate,
T_r – random repair time,
$G(t)$ – p.d.f. for repair time; $\overline{G}(t) = 1 - G(t)$,
λ^s – failure rate of new unit,
λ^n – failure rate of repaired unit.

The solution of the presented model is obtained with the use of the graphical method based on the Lorenz transform.

The main classification of the models regarding the optimisation of inventory and maintenance policy parameters is presented in [11].

3. An example

Consider a repairable system of systems under continuous monitoring, in which there are integrated two independent systems: operational and its supporting system. Both systems have only two states: upstate, when they are operable and can perform its specified functions, and downstate, otherwise.

The system of systems total task is defined as the continuous performing of exploitation process. Moreover, in the presented model, the logistic support functions are narrowed down only to providing the necessary spare parts to the operational system. As a result, the logistic support system is inoperable when there is no capability of supplying the operational processes with necessary spare parts.

The operational system is composed of M identical elements working in a reliability structure, which determines the moments when the system goes to a downstate. Let us also assume that element failures are random in time, and each failure entails a random duration of repair before the element/system is put back into service. Let us also assume that any information about failures in this system is reliable and comes immediately to the logistic system.

In the investigated model, when logistic support system is in upstate, the ability of the system of systems depends only on the following:

- The time, when the operational system is operable, and
- The time of technical system repair.

In this situation, when the supporting system is inoperable due to the lack of spare parts, the system of systems availability also depends on the logistic delayed time, which is necessary to solve logistic problems.

Moreover, if there is a restriction on the system of systems' total task completion time, defined as the time of operational system recovery process, the system of systems remains in upstate if this defined time will be shorter than time resource. Otherwise, the system of systems will fail and remain in downstate until the end of the delivery process.

Consequently, the following additional model assumptions are taken into account to define the system of systems performance process:

- The randomness and independence of all the performed processes,
- Critical inventory level (CIL) used as a stock policy, and
- the individual time redundancy used to model the system of systems performance [11].

To the best of the authors' knowledge, an effective way for achieving the reliable operational systems logistic support is especially based on meeting two targets: reliability/availability and cost constraints [11]:

$$\begin{cases} C_s = \dfrac{C_{sj}}{E[T_j]} \\ A \le A_{\min} \end{cases} \tag{16}$$

Where: C_{sj} – the expected total system of systems costs in a *j*th procurement cycle,
T_j – the random time of the *j*th cycle,
A – the system of systems' availability ratio,
A_{min} – the limiting availability ratio of system of systems performance.

More information can be found in [11].

3.2. Simulation model and obtained results

The analytical model of the performance of the presented system of systems with time dependency is investigated in, e.g. [16, 17, 18].

The analytical results of the modelled problem can be obtained only for a small amount of cases, when the operational system is a single-unit system and the performed processes are modelled according to the exponential distributions, *etc*. (see. [11]). Thus, there can be written the following conclusion, that this analytical model is an oversimplified version of the real system behaviour, so the obtained results are not traceable to practical situations.

To overcome this problem, there is proposed a simulation model of time dependent system of systems performance, which has been developed with the use of *GNU Octave* program. The simulation algorithm of the modelled system of systems is given in Fig. 4.

The systems of systems level of availability ratio, the probability of system of systems downtime occurrence, or economic results strongly depend on the operational system reliability structure. That is why the model of the time dependent system of systems performance was created for the three various system reliability structures – series, parallel, and "*k* out of *n*."

The simulation results of the modelled system of systems have been carried out for the input parameters, presented in Table 1.

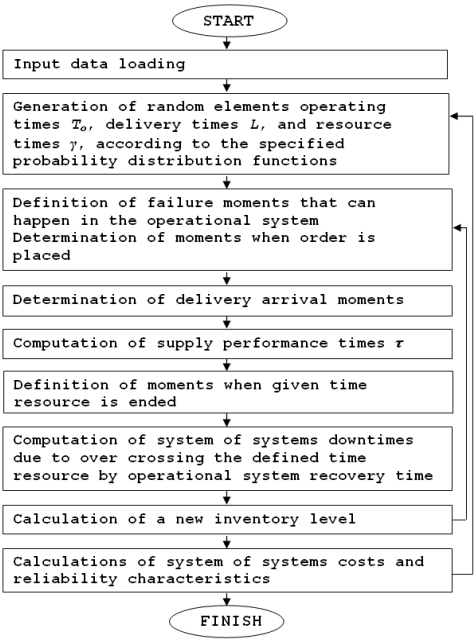


Fig. 4. Simulation algorithm of time dependent system of systems performance [11]
Rys. 4. Algorytm realizacji komputerowej symulacji procesu funkcjonowania nadsystemu [11]

Table 1. Input parameters of modelled system of systems [11]
Tabela 1. Parametry wejściowe modelu nadsystemu [11]

	Initial value	Explanation of denotation
A_o	1	Weibull's shape parameter of single operational element time to failure
$1/B_o$	100	Weibull's scale parameter of single operational element time to failure
A_r	1	Weibull's shape parameter of single operational element replacement time
$1/B_r$	10	Weibull's scale parameter of single operational element replacement time
A_L	1	Weibull's shape parameter of lead-time time
$1/B_L$	1000	Weibull's scale parameter of lead-time time
A_γ	1	Weibull's shape parameter of time resource time
$1/B_\gamma$	10	Weibull's scale parameter of time resource time
k	1	"k" out of "M"
M	5	number of elements working in an operational system
s	5	critical inventory level
Q	30	order quantity
c_w	50	replacement cost of a unit
c_o	50	ordering cost
c_p	50	purchase cost of one unit
c_h	100	inventory unit holding cost per unit time
k_f	1000	penalty cost of system of systems failure occurrence
c_{dw}	1000	cost of system of systems downtime unit

The main reliability and economic results are presented in Fig. 5–10.

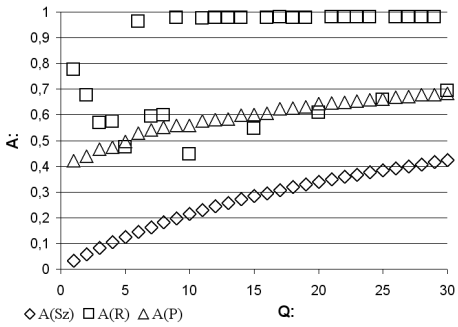


Fig. 5. System of systems availability ratio for various levels of order quantity of spare elements
Rys. 5. Zmienność poziomu wskaźnika gotowości w funkcji zmiany parametru wielkości partii zamówienia

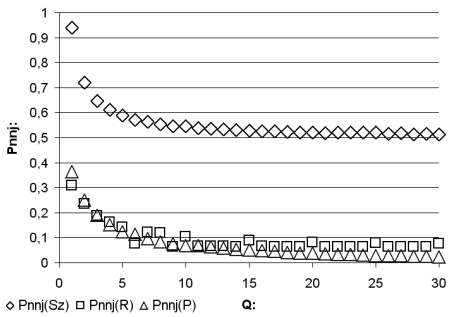


Fig. 6. The system of systems downtime probability for various levels of order quantity of spare elements
Rys. 6. Zmienność poziomu prawdopodobieństwa niezdatności nadsystemu w funkcji zmiany parametru wielkości partii zamówienia

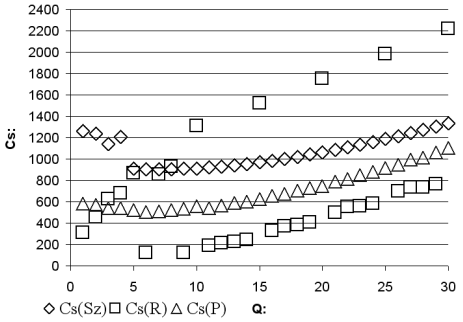


Fig. 7. Expected cost per unit time function for various levels of order quantity of spare elements
Rys. 7. Zmienność oczekiwanych kosztów funkcjonowania nadsystemu w funkcji zmiany parametru poziomu wielkości zamówienia

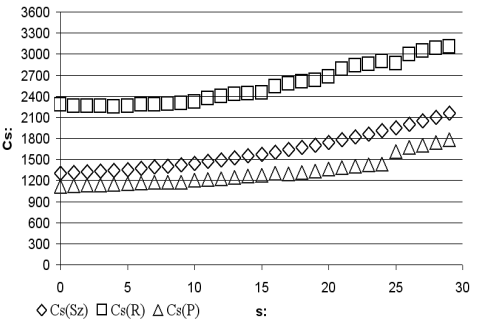


Fig. 8. Expected cost per unit time function for various critical inventory levels
Rys. 8. Zmienność oczekiwanych kosztów funkcjonowania nadsystemu w funkcji zmiany parametru poziomu zapasu alarmowego

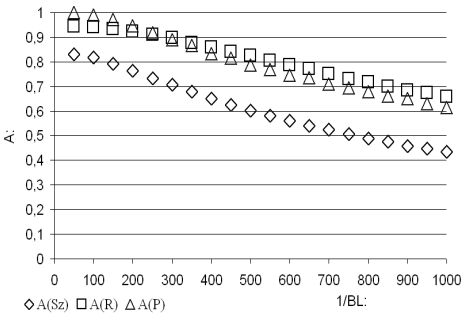


Fig. 9. System of systems availability ratio for various lead-time lengths
Rys. 9. Zmienność poziomu wskaźnika gotowości w funkcji zmiany parametru oczekiwanego czasu realizacji dostaw

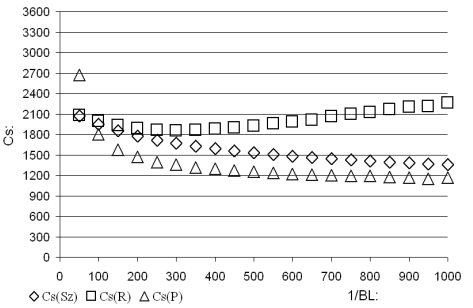


Fig. 10. Expected cost per unit time function for various lead-time lengths
Rys. 10. Zmienność oczekiwanych kosztów funkcjonowania nadsystemu w funkcji zmiany parametru oczekiwanego czasu realizacji dostaw

The ordered and delivered spare parts quantity determines the length of a single procurement cycle (time that elapses between the two consecutive moments when the inventories on-hand drop to a critical level). As a result, the bigger the ordered quantity, the higher mean stock level in the system and rarer deliveries are performed.

The expected costs incurred by the system of systems with a different-structured operational system are mainly determined by the inventory holding costs and the system of systems downtime costs (Fig. 7). When the level of ordered quantity rises, the expected costs function has the local minimum in the case of series and k out of M systems. It is a result of rarer deliveries and downtimes occurrence that arise from inventory lack. On the other hand, the

more spare elements are purchased, the higher inventory holding costs are incurred.

The worst solutions for the system of systems with an operational system in parallel occur when the ordered quantity Q is a multiple of M. If all M elements are replaced and there is no spare parts remaining in a logistic system, there is a higher downtime probability and economical consequences, than if there are some elements in a stock. These downtime costs together with the inventory holding costs have the greatest influence on the system of systems economic results.

The same effect can be seen when availability of the system of systems with the operational system in parallel is analysed. The system of systems reaches the lowest availability ratio level when Q is a multiple of M. For the system of systems with other reliability structures of operational systems, the rarer and bigger deliveries, the higher availability ratio is achieved.

The level of ordered quantity also has an influence on the probability of the system of systems downtime occurrence, which is especially evident for a series structure case (Fig. 6). A lower level of ordered quantity forces frequent deliveries, and as a result, there is higher probability that the possible delays of the delivery cause the system of systems downtime.

The next parameter of the procurement process, which affects the system of systems performance, is the critical inventory level (Fig. 8). The higher critical inventory level, the higher the mean inventory level in a system what incurs higher inventory holding costs. However, the higher s level also gives the possibility to reduce the delivery delay consequences, which has a positive impact on the system of systems reliability results.

Moreover, there can also be seen the influence of lead-time length on the system of systems behaviour. The longer lead-time especially affects the reliability characteristic of the system of systems (Fig. 9). On the other hand, the longer the lead-time, the lower inventory holding costs and the higher downtime costs incurred, which is connected with the higher probability of the system of systems downtime occurrence. This relation can be seen in the Fig. 10 as a local minimum of the C_s value.

In order to model the time dependencies between the operational system and its logistic support system, basic relations have to be identified, which result from the system of systems structure, component parameters, or process execution times.

In other words, the presented model can especially support decision processes in the area of supply task performance requirements. It especially gives a convenient tool to decide which supplier can provide the desirable time of supply delivery in order to achieve a defined system of systems operational capability.

On the other side, the developed model can be helpful to assess the reliability requirements of operational system elements in order to provide the continuous system of systems total task performance.

Conclusions

To sum up, all the presented models from the area of supply process parameters optimisation, when a system is maintained according to defined PM, can be divided into two groups:

- Searching for effective optimisation methods for already known models (see e.g. [37, 38]), and
- Searching for system models in which new assumptions are made (e.g. new reliability structure, dependent elements in a system) (see e.g. [23, 24]).

Moreover, literature on modelling relations between logistic and operational systems is scarce. Up to now, the interactions between the operational system and its supporting systems have not been clearly investigated. The research has focused on the evaluation of reliability and economical characteristics for both systems in a separate way.

Moreover, the logistic systems have been evaluating and designing mostly in terms of inventory modelling, supply processes organisation, and transportation processes modelling.

However, the simultaneous setting of all structural parameters (e.g. redundancy, repair shop capacity) and control variables (e.g. spare part inventory levels, maintenance policy parameters, repair job priorities, time resource) is mathematically a hard problem, and cannot be done without many simplified assumptions taken.

References

[1] Pierskalla, W.P. & Voelker, J.A., A survey of maintenance models: the control and surveillance of deteriorating systems. Naval Research Logistics Quarterly 23, (1976) 353-388

[2] Valdez-Flores, C. & Feldman, R., A survey of preventive maintenance models for stochastically deteriorating single-unit systems. Naval Research Logistics Vol. 36, (1989) 419-446

[3] Cho, I.D. & Parlar, M. , A survey of maintenance models for multi-unit systems. European Journal of Operational Research 51, (1991), 1-23

[4] Nakagawa, T., A summary of discrete replacement policies. European Journal of Operational Research 17, (1984) 382-392

[5] Nicolai, R.P. & Dekker, R., Optimal maintenance of multicomponent systems: a review. Economic Institute Report 2006

[6] Nowakowski T. & Werbińska S., Maintenance modelling concepts – state of art. International Journal of Materials and Structural Reliability. vol. 6, nr 2,(2008) s. 229-254

[7] Pham, H. & Wang, H., Imperfect maintenance. European Journal of Operational Research 94, (1996) 425-438
[8] Sherif, Y.S., Reliability analysis: Optimal inspection & maintenance schedules of failing equipment. Microelectronics and Reliability Vol. 22, No. 1, (1982) 59-115
[9] Thomas, L.C., A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. Reliability Engineering 16, (1986) 297-309
[10] Wang, H., A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 139, (2002) 469-489
[11] Werbińska S., Model of logistic support for exploitation system of means of transport. PhD thesis (not publ.). Wroclaw Technical University, Wroclaw, (2008)
[12] Sarjusz-Wolski, Z., Strategy of supply management. Practice of business logistics. (in Polish). A.W. PLACET, Warsaw, (1998)
[13] Guide Jr V.D.R., Srivastava R., Repairable inventory theory: Models and applications, European Journal of Operational Research, 102, (1997)
[14] Kennedy W.J., Patterson J.W. & Fredendall L.D., An overview of recent literature on spare parts inventories, International Journal of Production Economics, 76, (2002)
[15] Crossley W.A.: System of systems: an introduction of Purdue University Schools of Engineering Signature Area. Available on: http://esd.mit.edu/symposium/pdfs/papers/crossley.pdf (29.11.2006)
[16] Werbińska S., Interactions between logistic and operational system – an availability model, Risk, reliability and societal safety. Eds T. Aven, J. E. Vinnem, Leiden: Taylor and Francis, (2007)
[17] Werbińska S., The reliability model of logistic support system with Time Dependency (in Polish). Logistics No. 3, (2007)
[18] Werbińska S., Problems of designing logistic support system with time dependency. SYSTEMS Journal of Transdisciplinary Systems Science, Vol. 12, No. 2, (2007)
[19] Diaz A., Fu M.C., Multi-echelon models for repairable items: a review, Working Paper, University of Maryland, (1995)
[20] Diks E.B., de Kok A.G., Lagodimos A.G., Multi-echelon systems: A service measure perspective, European Journal of Operational Research, 95, (1996)
[21] Sridharan V., A note on ordering policies with cost of observation and lead times, International Journal of Production Economics, 32, (1993)
[22] Thomas L.C., Osaki S., An optimal ordering policy for a spare unit with lead time, European Journal of Operational Research, 2, (1978)
[23] Dohi T., Kaio N., Osaki S., On the optimal ordering policies in maintenance theory – survey and applications, Applied Stochastic Models and Data Analysis, 14, (1998)
[24] Dohi T., Osaki S., Kaio N., Optimal planned maintenance with salvage cost for a two-unit standby redundant system, Microelectronics and Reliability, Vol. 36, No. 10, (1996)
[25] Vujosevic M., Petrovic R., Senborn A., Spare parts inventory planning for a redundant system subject to a phased mission, Engineering Costs and Production Economics, 19, (1990)
[26] Chelbi A., Ait-Kadi D., Spare provisioning strategy for preventively replaced systems subjected to random failure, International Journal of Production Economics, 74, (2001)
[27] Hagmark P.E., Pernu H., Risk evaluation of a spare part stock by stochastic simulation, Safety and reliability for managing risk. Eds C. G. Soares, E. Zio, Leiden: Taylor and Francis, (2006)
[28] Sarker R., Haque A., Optimization of maintenance and spare provisioning policy using simulation, Applied Mathematical Modelling, 24, (2000)
[29] Mehrez A., Stulman A., Age replacement in the presence of inventory constraints, European Journal of Operational Research, 12, (1983)
[30] Vaughan T.S., Failure replacement and preventive maintenance spare parts ordering policy, European Journal of Operational Research, 161, (2005)

[31] Gharbi A., Kenne J-P., Beit M. , Optimal safety stocks and preventive maintenance periods in unreliable manufacturing systems, International Journal of Production Economics, 107, (2007)
[32] Kenne J-P., Gharbi A., Beit M., Age-dependent production planning and maintenance strategies in unreliable manufacturing systems with lost sale, European Journal of Operational Research, 178, (2007)
[33] Smith Ch. H., Schaefer M.K., Optimal inventories for repairable redundant systems with aging components, Journal of Operations Management, Vol. 5, No. 3, (1985)
[34] Park Y.T., Park K.S., Optimal stocking for replacement with minimal repair, Microelectronics and Reliability, Vol. 25, No. 1 (1985)
[35] Sheu S-H., Liou Ch-T., Tseng B-Ch., Optimal ordering policies and optimal number of minimal repairs before replacement, Microelectronics and Reliability, Vol. 32, No. 7, (1992)
[36] Dohi T., Ashioka A., Kaio N., Osaki S., The optimal repair-time limit replacement policy with imperfect repair: Lorenz Transform approach, Mathematical and Computer Modelling, 38, (20003)
[37] Catuneanu V.M., Moldovan C., Popentiu Fl., Gheorghiu M. , Optimum system availability and spare allocation, Microelectronics and Reliability, Vol. 28, No. 3. (1988)
[38] Migdalski J. (ed.), Reliability handbook. Mathematical basis. (in Polish) WEMA, Warszawa 1982

Manuscript received by Editorial Board, February11th, 2011

Modele utrzymania systemu technicznego – rola wsparcia logistycznego, przegląd

S t r e s z c z e n i e

W artykule przedstawiono przegląd literatury z zakresu modelowania logistycznego wsparcia funkcjonowania systemu technicznego ze szczególnym uwzględnieniem modeli zaopatrzenia systemów technicznych w części wymienne. Autorzy przedstawili klasyfikację danych modeli, wyróżniając cztery podstawowe grupy: modele doboru optymalnych parametrów strategii sterowania zapasami przy ustalonej strategii obsługi profilaktycznej obiektu, modele doboru optymalnych strategii sterowania zapasami zapewniających maksymalną niezawodność obiektu, modele niezawodności magazynowanych elementów wymiennych oraz modele doboru optymalnej strategii sterowania zapasami obiektu technicznego z wielopoziomowym systemem obsługiwania. W artykule skupiono się na omówieniu pierwszej z wymienionych grup modeli. Następnie w artykule przedstawiono model nadsystemu z rezerwą czasową. Wykorzystano analizę komputerowej symulacji w celu oceny wpływu parametrów procesu zaopatrzenia (wielkości partii zamówienia, poziomu zamawiania, czasu dostawy) na zachowanie nadsystemu.

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

JAN SZYBKA*

Methodology for reliability estimation of systems with sliding reserve

Key words

Reliability, sliding reserve, events, reliability structure of type k out of n.

Słowa kluczowe

Niezawodność, przesuwająca się rezerwa, zdarzenie, struktura niezawodnościowa typu k z n.

Summary

The main goal for using slipping reserve in renewable machine systems is to ensure high functional reliability with a simultaneous decrease to the minimum number of reserve elements.

The development of a theoretical basis for the estimation of the exploitation reliability of systems with sliding reserve was confined, in general, for the case of one reserve object that ensures the reliable function of a selected group of basic devices.

1. Introduction

Finding a solution to the problem of reserving for renewable technical devices with optional probability distributions for proper work is difficult from the theoretical point of view. The case under consideration refers to a class of processes broader than homogeneous processes. We consider independent

* Faculty of Mechanical Engineering and Robotics, University of Science and Technogy, 30-059 Kraków, Al. Mickiewicza 30 (tel. 12 617 31 10; e-mail: szybja@agh.edu.pl).

increments as streams of failures and renewals as streams without reciprocal interactions. Moreover, we assume that the parameters for the stream of failures, the stream of renewals, and the parameters for the failure of objects (which are in reserve) do not have to be constant. If those parameters are constant, then the behaviour of the system in defined by time intervals, and the probability of proper work can be determined on the basis of a system of differential equations [1, 2].

The assumption, which is often made, that the time for renewals has a constant value or is a random variable with an exponential probability distribution is not significant in this model. Very often, in practice, an assumption that streams of failures and renewals are straight streams (λ (t) = const.) is not correct. This may happen when (in spite of random factors which disturb proper work of the system) failures caused by the natural wear of elements of machines and devices start to dominate. An assumption that probability distributions for failures and renewals have optional shapes makes analytical solutions complicated. Although these problems can be solved by inserted Markow chains or methods based on differential-integral equations or Spicer identity, the results obtained in the form of complicated dependencies are often not practically useful [2].

2. Methodological assumptions

The methodological approach to the problem of reserving differs from traditional approaches in which problems with reserving are solved by adjustment of the number of reserve devices to the number of working basic devices. If the optimum number of basic devices is known and the reliable function of these devices is ensured by a single reserve object by multiplication of the calculated system, it is possible to increase the number of reserve objects for more numerous groups of technical devices. Therefore, it is necessary to search for solutions to such problems by means of simulation experiments that can be carried out in continuous or discrete times [8].

The methodological approach, which is proposed, substantially facilitates (in comparison with methods used so far) making simulation models with high practical importance [2].

The object under study is a machine system, which consists of n elements. The reliability structure of the system is a threshold structure of type k out of n [4, 5], provided that k=n-1. All elements are renewable. Elements which are marked by code numbers from 1 to k are called basic elements; whereas, the element marked by subscript n is a renewable reserve element [6, 7, 8].

A reserve element plays a role of the “sliding reserve”. Term *element* can refer to either a technical object or a part of a machine. The meaning of this

word depends on the degree of generalisation made and does not play a significant role in this analysis.

The presented machine system has an excessive and coherent reliability structure described by an efficiency function, which, in turn, is a logical function of the following form:

$$F(X) = 1 \text{ for } X: f(X) \geq k,$$

$$F(X) = 0 \text{ for } X: f(X) < k.$$

The system is in an "able-to-work state" (aptitude) when F(X)=1, or in an "unable-to-work state" (non-aptitude) when F(X) = 0. It is preliminarily assumed that the introduction of the reserve device into work starts with reliability equal to 1. This assumption can be changed (weaken) if we assume that reliability is different from 1; however, this does not have a crucial significance in methodological analysis of the problem.

By the analysis of work of studied system, the following hypothetical reliability and exploitation states for system aptitude can be distinguished.

S_l – hypothetical reliability-exploitation states of technical system; l = 1,.....m, for basic elements that reside in states of work, reserve, and renewal [2].

A given object has an universal character and, on this basis, it is possible to examine other cases of reserving for systems that encompass, e.g., multiple slipping reserve when reliability of switches is taken into account or the application of this type of redundancies in complex, parallel-serial reliability structures. This model can also be used for forecasting the number of exchangeable parts, which can be considered as redundancies that belong to cold reserve [3, 9].

The concept of simulation models that are presented below is based on the *transition module* [3]. Analysis of transitions in the network of events leads to the assessment of the reliability of function for the whole system. In order to characterise of the simulation model, it is necessary to present terms and concepts formulated in the work [9].

The most important terms and concepts are the following:

– Reliability-exploitation states,
– Elementary events,
– Events,
– Relations between elements within the set of events, and
– Trajectories of events.

e_j – (code numbers for system elements) residence time in reliability-exploitation states for system elements.

They are briefly characterised below [2].

Ad a) It can be assumed that the physical state of an object [8] is characterised by measurable or non-measurable features, which are specified according to the needs of the analysis. Taking into account the ability of performing operational tasks by the object, it is possible to talk about its ability-to-work or inability-to-work. However, it is equally important to determine its degree of readiness for performing these tasks when the object being repaired and when it can start working or when it immediately starts to work in the able-to-work state. As a result, it was assumed the following:

Reliability state – refers to able-to-work state or unable-to-work state,

Exploitation state – refers to state of work, state of waiting for work (state of reserve) or state of renewal (which takes into account the waiting time for renewal).

Reliability and exploitation states are jointly examined, and they are described as **reliability-exploitation states**. This means that an element at any moment is in one of reliability states (aptitude or non-aptitude) and simultaneously in one of the exploitation states (work, reserve, or renewal).

Ad b) elementary event ($e_{ij?}$) – this term is exclusively associated with the exploitation of elements of the system and describes the shift of every element to the next moment of exploitation t (exploitation time is examined in a discrete way; the term *exploitation time* t is defined in work [8]. A set of elementary events that can occur is marked as follows:

$$\Omega = \{ \mathbf{e_{ijt}} : i = 1, 2, \dots, k, n ; j = 1, 2, 3 ; t = 1, 2, 3, \dots \}$$

i – number elements,
j – state of work (1), reserve (2) or renewal (3),
t – moment of exploitation.

The ability to work for the whole system depends on abilities of its elements. One non-aptitude element does not cause failure of the whole system; however, failure of two elements makes the whole system unable to fulfil operational tasks.

Ad c) Event ($E_{w,t}$) referring to changes which occur during exploitation of the whole system describes any shift of the system to the next moment in exploitation t. Occurrence of an event is identified with a change in operation properties of the system in subsequent moments of exploitation [6, 8], which depends on the reliability-exploitation states of elements.

The ability of the system to function can be described by using the above mentioned elementary events and events.

In order to do that, the set of elementary events $\{e_{ij?}\}$ should be transformed into a set of events $\{E_{w,t}\}$ which characterises the behaviour of the whole system within the time of exploitation:

$$\Psi: \{e_{ijt}\} \xrightarrow{\Psi} \{E_{w,t}\}$$

t – index which refers to exploitation time when the system was in the specified reliability-exploitation state

$$t = (1,2,.....),$$

w – index which describes the reliability-exploitation state of the analysed technical system.

This set of events $\{E_{w,t}\}$ represents hypothetical reliability-exploitation states of system's ability-to-work, that is, whether the operating system has able-to-work (aptitude) reserve or whether it operates with an additional unable-to-work (non-aptitude) element which is being repaired or whether the whole system failed because at least two elements were damaged.

The probability for proper work of the system with a reserve aptitude element is higher than that with an additional element that is being repaired. Therefore, distinguishing between these two reliability-exploitation states has substantial significance for reliability analysis.

Ad d) Systematisation of a set of events $\{E_{w,t}\}$ is done on the base of the following binary relations [9]: simultaneity and preceding relations for events of concurrent events.

The above mentioned terms, elementary event ($e_{ij?}$), event ($E_{w,t}$) and reliability-exploitation state (S_l), successfully describe the actual situation in a quite detailed way. This approach makes the building of theoretical models very complicated; therefore, this approach to the problem was simplified. In order to do this, specified relations of preceding, simultaneity, similarity, and affinity are used.

In this model, a significant role is played by subsets of events that are connected to each other by **relations of simultaneity (J), preceding (C) and affinity of events (K).** Some of them are briefly characterised below.

At any moment of exploitation (t), the technical system is in one of specified reliability-exploitation states, i.e., one of the events specified by the simultaneity relation occurs. In trajectories of events under evaluation, only events that belong to different classes distinguished by means of simultaneity relation (J) can precede each other.

The trajectory encompasses events that belong to different classes $\{A_i\}$ [9]. These transitions have a certain probability, for instance, the transition from the class of concurrent events A_1 to the class of concurrent events A_2.

Affinity relation of events (K) plays a significant role in developing the simulation model.

Subsets of the set of affinity events cover events that occur one by one in subsequent moments of exploitation. Subsets of affinity events consist of events that occur one by one in subsequent moments of exploitation and belong to the same trajectories. Set Z_p is a one-component set. Specified events precede each

other. Subsets of affinity events E_o and E_p are defined as reliability-exploitation states, S_1, S_2, ..., S_m, in which the system exists in subsequent moments of exploitation.

It is said that the event belongs to the set of events E_o when any of the elements of the technical system, but only one, undergoes renewal. It is said that the event belongs to the set of events E_r when any of the elements of the technical system, but only one, is in reserve. Any other event, which does not belong to set $E_r \cup E_o$ means that this element belongs to set E_p.

Events that belong to classes Er and E0 occur one by one in subsequent events. Each of these classes can be considered as a multiplet in an aspect of potential opportunities for the realisation of operational tasks (function) by the examined technical system. These events constitute a phase space through which proceeds the trajectory of a random process.

The number of affinity events that subsequently occur in the trajectory is associated with time needed for change in the reliability-exploitation state.

Subsequent time intervals of waiting for changes in the state of the trajectory of affinity events are non-negative random variables, which can have identical or different probability distributions during the exploitation of the system. Residence time in which the system resides in specified reliability-exploitation states and the probability of transitions to subsequent states (the occurrence of new events) are interesting from a practical point of view.

Therefore, the function that describes the probability of proper work is as follows:

$$R(t) = P(T > t).$$

In this work, reliability is concerned with the probability of keeping specified characteristics or parameters that characterised the proper work of the technical system in certain exploitation conditions and in a specified time interval.

Application of statistical studies and the theory of probability in order to assess the behaviour of the object during exploitation time is justifiable when there are problems with identification of the interaction of external factors (either the identification of what sort of factors influence correct function of objects and the strength of their interactions). This case often occurs in practice. The basic concept of reliability theory is term failure. Failure is treated as a partial or full loss of these properties of the object (technical system), which constitute the condition for proper function, according to a specified goal. According to probability theory, every complex reality can be described in categories of the occurrence or non-occurrence of certain events. By examining this reality within a specified time interval, we can assess, in statistical way, of the probability of the occurrence of certain events in subsequent moments of the exploitation of the event system [2].

Ad e) Behaviour of the technical system during exploitation can be described by means of trajectories $\{T_{zwt}\}$ defined in set $\{E_{w,t}\}$. The number of possible occurring trajectories of events is doubled in any subsequent moment of exploitation.

With reference to the goal of the study, it is important to distinguish such trajectories that characterise the system behaviour in a certain time interval. Because the trajectory consists of subsequent events, the actual problem is how to specify criteria for the selection and choice of preferred events in the analysed moment of exploitation.

If event $E_{0,t}$ occurs at moment t, it means that in subsequent moment t+1, event $E_{0,t+1,}$ or $E_{r,1}$ or $E_{p,1}$ can occur. Which event occurs depends on what principle for event selection is used. The set of events $\{E_{w,t}\}$ with trajectories $\{T_{zw,t}\}$ that were specified, according to accepted selection principle, constitutes the event system of exploitation. The selection principle can be identified with a relation specified on sets of events. The goal for the operation of a specified event system exists and is included within this system, because the set of events $\{E_{w,t}\}$ is specified according to this target. In addition, trajectories of events are considered according to this goal. It was assumed in the model described in this publication that the analysis of the processes of failures and renewals is performed in discrete time because, in this way, it is possible to use modelling based on the theory of events.

Occurrence of subsequent events can happen in a random or deterministic way. Time intervals between events can differ on the bases of length or they can be the same. They may have a random or deterministic character. The concept which is adopted in this publication is based on events of which occurrence specifically depends on “exploitation moments.” Digitisation of time, i.e. the analysis of the occurrence of subsequent events in subsequent moments of exploitation has a stationary character (time intervals are constant) and depends on actual exploitation conditions as well as on the character of the simulation experiment carried out. Residence time for the technical system in the same reliability–exploitation state is a random variable [8].

Subsequent events occur according to the random selection of transition in the “transition modules.” It is assumed that each reliability and exploitation state described by a set of relevant events, in which the system occurs, have transition characters. Possible transition to other states is determined by a relevant transition function. Residence time of the system in these states and trajectories of transitions in the space of events determine the dynamics of changes which happen in the system. The concept of transition modules makes it possible to determine probability values for the residence of the system in an able-to-work state at any moment of exploitation. The system is in an able-to-work state if basic elements perform production tasks and reserve element n is in repair (e.g. state S_1) or has just been repaired (e.g. state S_2) and is waiting to be introduced into work.

The idea of building of transition modules is illustrated on the base of the analysis of the ability-to-work state of the system as presented. It is possible to distinguish the following cases according to the results of the assessment of the able-to-work state of the system at moments from t_1 until t+1:

- At moment t, all elements of the system from 1 to k perform operation tasks and reserve element n is being repaired ($E_{0,t}$). Within a short time interval, which follows moment t through moment t+1, the situation does not change ($E_{0,t+1}$); or,
- In moments from t to t+1, all elements of the system from 1 to k work and reserve element n which was in repair at moment t was repaired within the time interval from t to t+1 and is waiting to be introduced into work ($E_{r,1}$); or,
- At moment t+1, a failure of the system ($E_{p,1}$) happens. This can happen when element n within time interval t to t+1 was not repaired and, in the meantime, another working element was damaged.

Failure of the whole technical system always occurs when an event that belongs to the class of events Ep happens. Reliability analysis of the function of the technical system should start with output event, i.e., event $E_{0,1}$.

Linearly ordered moments determine exploitation time [8] and the above mentioned events occur in stroke mode. Residence times for the system in relevant reliability-exploitation states, which are counted by means of the number of transitions (elements) of the system through subsequent exploitation moments t without changes in reliability-exploitation states are representative attributes of the quality of the whole system and its elements (physical technical objects). Indices for shares of the able-to-work state, work and renewal states, with reference to the global time of exploitation, are particularly important in system reliability assessment.

On the base of assumptions made in work [8], which were partially characterised above, principles for the analysis of transition modules and making of the simulation model are given in further parts of this work.

3. Conclusions

The concept of modelling the reliability of systems with sliding reserve type k out of n is described in detail in [2, 8]. The idea we have presented in this paper has been applied in simulation the modelling of reliability. This article is the first in a series of the forthcoming papers about this topic. The forthcoming articles will concern simulation methods of reliability estimation. The approach we have presented is unique, and it has not been proposed in literature before (except the articles of the author of this publication). It may be successfully applied to the reliability problems together with standard analytical methods.

References

[1] Bobrowski D.: Probabilistyka w zastosowaniach technicznych. Warszawa, WNT, Warszawa 1985.
[2] Szybka J.: Prognozowanie niezawodności urządzeń mechanicznych funkcjonujących w układach z rezerwą. Wydawnictwa AGH, z. 34, Kraków 1996.
[3] Barlow R.S., Heidtman K.D.: System Reliability. IEEE Trans. Reliability, Oct. 1984, Vol. R-33, (322-333).
[4] Biernat J.: An Iterative Algoritm for Computing the Reliability of Computing k-out-of-n. PAN – Zagadnienia Eksploatacji Maszyn z.3 (93). 1992.
[5] Jain S.P., Gopal K.: Recursive Algorithm for Reliability Evaluation Computing k-out-of-n : G Sysytem. IEEE Transaction on Reliability, Vol. R-34, June 1985.
[6] Szybka J.: Zagadnienia rezerwowania w eksploatacji odnawialnych systemów maszynowych z rezerwą przesuwającą się. Zeszyty Naukowe AGH, Mechanika nr 16, Kraków 1989.
[7] Betzin O., Runge W.: Zastosowanie metod matematycznych do wyznaczania rezerw awaryjnych. PNIOiZ Polit. Wrocławskiej nr 26, Wrocław 1976.
[8] Szybka J.: Zastosowanie teorii zdarzeń w modelowaniu eksploatacji. Zeszyty Naukowe AGH, Automatyka, z. 42, Ref. Na VII Ogólnopolskim Sympozjum – zastosowanie teorii systemów, Kraków 1987.
[9] Ważyńska-Fiok K., Jaźwiński J.: Niezawodność systemów technicznych. Warszawa, PWN, 1990.

Manuscript received by Editorial Board, February 14th, 2011

Metodyka szacowania niezawodności układów z przesuwającą się rezerwą

S t r e s z c z e n i e

W artykule przedstawiono metodyczne podstawy przeprowadzania oceny niezawodności układów o strukturze typu k z n. Tego typu układy są często stosowane w praktyce, ponieważ stanowią efektywny sposób zwiększenia niezawodności układów z rezerwą. Najskuteczniejszym sposobem tego typu rezerwowania jest układ podwajany, ale koszt jego stosowania jest najwyższy. Stąd też ocena, ile urządzeń pracujących powinno przypadać na jeden obiekt obiekt rezerwowy będący rezerwą przesuwającą się jest interesującym problemen z praktycznego i teoretycznego punktu widzenia.

W celu zbudowania modeli oceny niezawodności przeprowadzono analizę funkcjonowania układów z rezerwą przesuwającą się w ujęciu zdarzeniowym. Wyróżniono i zdefiniowano podstawowe pojęcia i opracowano sposób oszacowania niezawodności. Sformułowane założenia metodyczne umożliwiły zbudowanie modeli symulacyjnych szacowania niezawodności układów z rezerwa przesuwającą się, znajdujących zastosowanie do szybkiej oceny niezawodności różnych form rezerwowania.

Artykuł jest wprowadzeniem do serii opracowań na temat niezawodności układów typu k z n, z rezerwą przesuwającą się.

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

JAROSŁAW ZIÓŁKOWSKI*

Method of formulating the required number of tankers for delivery aircrafts in aviation fuel

Key words

Aircraft, tanker, supply, aviation fuel.

Słowa kluczowe

Statek powietrzny, cysterna, zaopatrywanie, paliwo lotnicze.

Summary

An important part of the air base logistic system is the supply sub-system. In military operations the main delivery can be focused on munitions and aviation fuel. Effective management of the supply stream and the reliability of vehicles in the air base logistic system affect the quality of operations, which can be measured by on time provisions, economic factors and the reliability of tankers. At present the number of tankers in the air base logistic system is based on experiences. This paper presents a mathematical model that enables one to estimate the minimum number of tankers in dependence from the type of the aircrafts, the number of aircrafts, the length of flights and structure of the flights.

1. Introduction

The main aim of the air base logistics system is meeting the needs of military technology. The following areas create the structured elements of that system:

* Military University of Technology, 00-908 Warsaw, Kaliskiego 2 St., phone. +48 22 6837109, e-mail: jziolkowski@wat.edu.pl

– Supply,
– Transportation,
– Infrastructure,
– Furnishing of services,
– Fire protection,
– Repair and serviceability, and,
– Airborne engineering supply.

The important elements of supplying aircrafts with aviation fuel are tankers. The extract number of them influences the punctuality of the flights and the quality of air force training and operational readiness. Nowadays this number is empirical. This article presents a mathematical model of estimating the required number of tankers.

2. Mathematical model of estimating the required number of tankers

The assumptions taken to formulate the mathematical model of supplying aircraft in flight with aviation fuel are as follows:

– The tanker can, be in one of specified operational states at any time.
– The number of aircrafts participating in flights is a random variable.
– The fuel consumption index of the main tank is a random variable and takes values $K_{zu} \in \{0.33; 0.5; 0.66; 0.85\}$.
– The storage tanks can be damaged at random during operation.
– The time needed to exchange a used storage tank for an operational one is specified.
– The duration of flights is fixed.

The following variables were taken to formulate the model:

- The number of aircrafts - N_{SP};
- The capacity of main tank of the aircrafts –V_{zbsp};
- The fuel consumption index of the aircraft– K_{zu};
- The number of storage tanks - N_{CD};
- The capacity of storage tanks – V_{CD};
- The flights structure (according to the planned schedule).

The fuel balance Equation for a single squadron of flights in accordance with the set time duration can be estimated as follows:

$$\sum_{k}^{N_{sp}} {}_{k}K_{zu} \cdot V_{zbsp} = V_{elt1} \tag{1}$$

Where: k = 1,2, ..., means the aircraft number.

The used fuel balance Equation for the maximum number of flights of a tactical aviation squadron can be set as follows:

$$\sum_{k}^{N_{SP}} \sum_{l}^{{}_{k}N_{elt}} ({}_{lk}K_{zu}) \cdot V_{zbsp} = V_{elt\max} \tag{2}$$

Where: ${}_{lk}N_{elt}$ – the l flight of the k squadron aircraft;
${}_{lk}K_{zu}$ – the fuel consumption index for l flight of the k aircraft;
$N_{elt} \in \{1,...,8\}$ – the number of flights.

Equation (2) is based on the assumption that all aircrafts where in flights.

If not, zero must be taken for the l flight of the k aircraft. The refuelling state of the aircrafts was considered for 2 situations:

1) for zero waiting time (t_{ocz}=0);

$$t_4 = t_m + K_{zu} \times t_{et} \tag{3}$$

Where: t_4 – the time of refuelling;
t_m – the handling time (connected with the time of approaching a new storage tank and connecting the injection sprayer);
K_{zu} – fuel consumption index;
t_{et} – the time of refuelling an empty tank of the aircraft.

2) $t_{ocz} \in$ (40min. - t_4) – for the maximum flight frequency (e.g. every 40 min).

For K_{zu}=0,85 and T_0=8h, we obtain for the squadron N_{elt}=5 and the maximum amount of the used fuel is:

$$V_{elt\max} = \sum_{k=1}^{16} \sum_{l=1}^{5} K_{zu} \cdot V_{zbsp} \tag{4}$$

The balance Equations of the capacity and work time for storage tanks are described with Equations (5) and (6) as follows:

a) The capacity balance:

$$N_{CDV} \cdot V_{CD} \geq V_{elt\max} \quad \text{Hence: } N_{CDV} \geq \frac{V_{elt\max}}{V_{CD}} \tag{5}$$

Where: N_{CDV} – the number of storage tanks required for the all fuel used by the squadron;
V_{CD} – the capacity of the storage tank;
V_{eltmax} – the maximum fuel consumption for the squadron while performing the task.

$$N_{CDV} \geq \begin{Bmatrix} 69.889 & dla & V_{CD-4,5} \\ 41.933 & dla & V_{CD-7,5} \end{Bmatrix}$$

b) The balance of the times: It was assumed that the time of the refuelling cycle for storage tank t_5 must fulfil the following condition:

$$t_5 \leq t_{nCDV} + t_m + t_2 + t_3 \; ; \tag{6}$$

Where: t_{nCDV} – the time of filling up the storage tank of a particular capacity dependent on its refuelling state;
t_m – the handling time, connected with approaching the depot;
t_2 – the required dwell time;
t_3 – the time of fuel quality inspection in a storage tank.

It was assumed, that the storage tank in T_0 could perform at most N_5 refuelling cycles. The Equation at the work time balance for one storage tank, which is needed for examining the feasibility of the process, can be defined as follows:

$$t_5 \cdot N_5 + t_4 \cdot N_4 \leq T_0 \tag{7}$$

$$N_5 \leq \frac{T_0}{t_5} \qquad\qquad N_4 \leq \frac{T_0}{t_4} \tag{8}$$

Where: t_5 – the time of refuelling cycle by a storage tank;
N_5 – the number of refuelling cycles;
t_4 – the time of refuelling of aircraft;
N_4 – the number of the aircraft refuelling;
T_0 – the flight duration.

If $V_{CD} \approx V_{zbsp}$, then for one storage tank $N_4 = N_5$ must be taken and the time margin appears. As a rule $t_4 \neq t_5$, but in practice $t_4 << t_5$.

The number of storage tank N_{CDV} needed by volume that provide the squadron which the all fuel must be converted to the number actually needed storage tanks N_{CD} according to the following Equation:

$$N_{CD} \geq \frac{N_{CDV}}{N_5} \tag{9}$$

Where: N_{CD} – the number of actually needed storage tanks for the flights;
N_{CDV} – the number of storage tanks needed by volume for the correct amount of fuel required in one flight;
N_5 – the number of possible refuelling cycles for a storage tank.

Based on these assumptions, we have the following Equations:

$$\sum_{k}^{N_{SP}} \sum_{l}^{{}_k N_{elt}} ({}_{lk}K_{zu}) \cdot V_{zbsp} = V_{elt\max} \tag{10}$$

$$N_{CDV} \geq \frac{V_{elt\max}}{V_{CD}} \tag{11}$$

$$N_{CD} \geq \frac{N_{CDV}}{N_5} \tag{12}$$

$$t_5 \cdot N_5 + t_4 \cdot N_4 \leq T_0 \tag{13}$$

$$N_5 \cdot V_{CD} \geq N_4 \cdot V_{zbsp} \cdot K_{zu} \tag{14}$$

We can then determine the range of feasible solutions on limited assumptions.

3. Numerical example

Basic assumption (variables):

- the number of aircrafts-8;
- the duration of the flights T_0 = 480 h.;
- the flights performed on Su-22 aircrafts;
- the fuel consumption index is dependent on flight duration (according to the graph 1) K_{zu}= {0.33; 0.5; 0.66);
- the flight structure according to the planned schedule (Graph. 1).

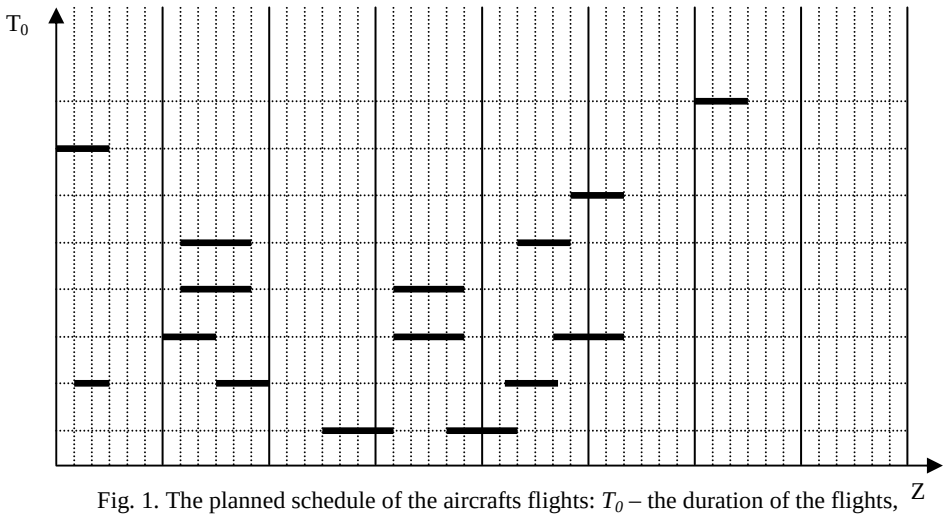


Fig. 1. The planned schedule of the aircrafts flights: T_0 – the duration of the flights, Z – the number of aircrafts
Rys. 1. Planowa tabela lotów – wariant: T_0 – czas wykonywania lotów, Z – liczba statków powietrznych

The minimum number of storage tanks (capacity 7.5 m^3) that will guarantee flight continuity must be specified.

The solution from: Equation (10) of the amount of fuel used by the aircraft in flights can be calculated as follows:

$$V_{elt} = \sum_{k}^{N_{SP}} \sum_{l}^{{}_{k}N_{elt}} ({}_{lk}K_{zu}) \cdot V_{zbsp} = 38991.25 \ [\mathrm{dm}^3]$$

Since the duration of the aircrafts is described with a discrete random variable for which the fuel consumption index takes values from K_{zu}= {0.33; 0.5; 0.66), the required number of storage tanks must be calculated for three cases and then the partial results summed.

Case I

The amount of the used fuel for K_{zu}= 0.66

$$V_{0,66} = 0{,}66\ (\mathbf{2}*V_{1zbsp}+2*V_{3zbsp}+2*V_{4zbsp}+1*V_{5zbsp}) = 21277.5\ [\mathrm{dm}^3]$$

The particular number of storage tanks which have a fixed capacity according to the Equation (11) must be bigger or the same as the needed amount of fuel:

$$N_{CDV} \cdot V_{CD} \geq V_{elt}$$

$$N_{CDV} \geq \frac{V_{elt}}{V_{CD}} \geq \frac{21277.5}{7500} \geq 2.83 = 3$$

The time of the refuelling cycle consist of the group of different times: The time needed for reaching the store, the handling time, the exact time of refuelling, the time of coming back to the surface of the airfield, the dwell time and quality control of the fuel. It is:

$$t_5 = 2.54 + 54 = 56.54\ [\mathrm{min}]$$

The refuelling time of the aircraft is precisely determined, because it depends on the fuel consumption index (K_{zu}), the efficiency of the tanker and the time needed to reach the aircraft by the tanker:

$$t_4 = 0.66*4625/300+5 = 15.17\ [\mathrm{min}]$$

The number of refuelling cycles both by storage tank *(N_5)* and the aircraft *(N_4)* can be calculated, according to Equations (13 and 14), by solving the following system of inequalities:

$$\begin{cases} t_5 \cdot N_5 + t_4 \cdot N_4 \leq T_0 \\ N_5 \cdot V_{CD} \geq N_4 \cdot V_{zbsp} \cdot K_{zu} \end{cases}$$

$$N_5 \le \frac{T_0}{t_5} - \frac{t_4 \cdot N_4}{t_5} \le \frac{480}{56.54} - \frac{15.17 \cdot N_4}{56.54} \le 8.48 - 0.268 N_4$$

$$N_5 \ge \frac{K_{zu} \cdot V_{zbsp} \cdot N_4}{V_{cd-7,5}} \ge \frac{0.66 \cdot 4625 \cdot N_4}{7500} \ge 0.407 N_4$$

$$8.48 - 0.268 N_4 = 0.407 N_4$$

$$0.675N_4=8.48 \;\Rightarrow\; N_4=12.56$$

$$N_5 \ge 0.407 N_4 \ge 5.11 = 6$$

From Equation (11) the number of storage tanks needed for one flight can be calculated as follows:

$$N_{CDV} \ge \frac{V_{elt}}{V_{CD}} \ge \frac{21277.5}{7500} \ge 2.83 = 3$$

After substituting N_{CDV} and N_5 (12) we get:

$$N_{CD0,66} \ge \frac{N_{CDV}}{N_5} \ge \frac{3}{6} = 0.5$$

Case II

The amount of the used fuel for K_{zu}= 0.5

$V_{0.5}$= 0.5 **(2***V_{2zbsp}+V_{3zbsp}+V_{5zbsp}+V_{6zbsp} + V_{7zbsp} + V_{8zbsp})= 16187.5 [dm^3]

The particular number of storage tanks of fixed capacity must be bigger or the same as the maximum amount of fuel. It can be calculated by the following Equation:

$$N_{CDV} \cdot V_{CD} \ge V_{elt\,\max}$$

$$N_{CDV} \ge \frac{V_{elt}}{V_{CD}} \ge \frac{16187.5}{7500} \ge 2.158 = 3$$

The time t_5 of refuelling cycle by a storage tank is as follows:

t_5= 1.927 +54 = 55.927 [min]

The refuelling time of the aircraft is precisely determined, because it depends on the capacity of the aircrafts tank, the fuel consumption index, the efficiency of the tanker and the time need to reach the aircraft by the storage tank:

t_4= 0.5*4625/300+5 =12.71 [min]

The number of refuelling cycles by the storage tank *(N_5)*, as well as the number of refuelling aircraft *(N_4)*, can be calculated from the following system of inequalities:

$$\begin{cases} t_5 \cdot N_5 + t_4 \cdot N_4 \le T_0 \\ N_5 \cdot V_{CD} \ge N_4 \cdot V_{zbsp} \cdot K_{zu} \end{cases}$$

$$N_5 \le \frac{T_0}{t_5} - \frac{t_4 \cdot N_4}{t_5} \le \frac{480}{55.927} - \frac{12.71 \cdot N_4}{55.927} \le 8.58 - 0.227 N_4$$

$$N_5 \ge \frac{K_{zu} \cdot V_{zbsp} \cdot N_4}{V_{cd-7.5}} \ge \frac{0.5 \cdot 4625 \cdot N_4}{7500} \ge 0.308 N_4$$

$$8.58 - 0.227 N_4 = 0.308 N_4$$

$$0.535 N_4 = 8.58 \quad \Rightarrow \quad N_4 = 16.03$$

$$N_5 \ge 0.308 N_4 \ge 4.93 = 5$$

From Equation (11) we calculate N_{CDV} :

$$N_{CDV} \ge \frac{V_{elt}}{V_{CD}} \ge \frac{16187.5}{7500} \ge 2.15 = 3$$

After substituting N_{CDV} and N_5 we get:

$$N_{CD0.5} \ge \frac{N_{CDV}}{N_5} \ge \frac{3}{5} = 0.6$$

Case III

The amount of the used fuel for K_{zu}= 0.33

$$V_{0.33} = 0.33\ (V_{2zbsp}) = 1526.25\ [dm^3]$$

The particular number of storage tanks of fixed capacity must be bigger or equal as the needed amount of fuel. It can be calculated as follows:

$$N_{CDV} \cdot V_{CD} \ge V_{elt\,\max}$$

$$N_{CDV} \ge \frac{V_{elt}}{V_{CD}} \ge \frac{1526.25}{7500} \ge 0.2 = 1$$

The time t_5 of refuelling cycle by a storage tank is as follows:

$$t_5 = 1.27 + 54 = 55.27\ [min]$$

The time t_4 of refuelling is precisely estimated and is as follows:

t_4= 0.33*4625/300+5 =10.08 [min]

The number of refuelling cycles by the storage tank *(N_5)*, as well as the number of refuelling aircraft *(N_4)*, can be calculated from the following system of inequalities:

$$\begin{cases} t_5 \cdot N_5 + t_4 \cdot N_4 \le T_0 \\ N_5 \cdot V_{CD} \ge N_4 \cdot V_{zbsp} \cdot K_{zu} \end{cases}$$

After substituting data we get the following:

$$N_5 \le \frac{T_0}{t_5} - \frac{t_4 \cdot N_4}{t_5} \le \frac{480}{55.27} - \frac{10.08 \cdot N_4}{55.27} \le 8.68 - 0.18N_4$$

$$N_5 \ge \frac{K_{zu} \cdot V_{zbsp} \cdot N_4}{V_{cd-7.5}} \ge \frac{0.33 \cdot 4625 \cdot N_4}{7500} \ge 0.2035N_4$$

$$8.68 - 0.18N_4 = 0.2935N_4$$

$$0.3835N_4 = 8.68 \Rightarrow N_4 = 22.63$$

$$N_5 \ge 0.2035N_4 \ge 4.605 = 5$$

From Equation (11) we calculate N_{CDV} as follows :

$$N_{CDV} \ge \frac{V_{elt}}{V_{CD}} \ge \frac{1526.25}{7500} \ge 0.2 = 1$$

After substituting N_{CDV} and N_5 we get the following:

$$N_{CD0.33} \ge \frac{N_{CDV}}{N_5} \ge \frac{1}{5} = 0.2$$

The precisely required number of storage tanks (capacity 7.5 m^3) was achieved as a sum of the three partial -results for the particular fuel-consumption indexes *(K_{zu})* according to the following equation:

$$N_{CD-7.5} > N_{CD0.33} + N_{CD0.5} + N_{CD0.66}$$

Because 0.5 +0.6 + 0.2= 1.3 Hence $N_{CD-7.5} > 1.3 \Rightarrow$ **2**

Based on (13) and (14), the relation between the storage tank capacity V_{CD} and the number of its refuelling cycles N_4 can be estimated. The results are illustrated by Graph 2.

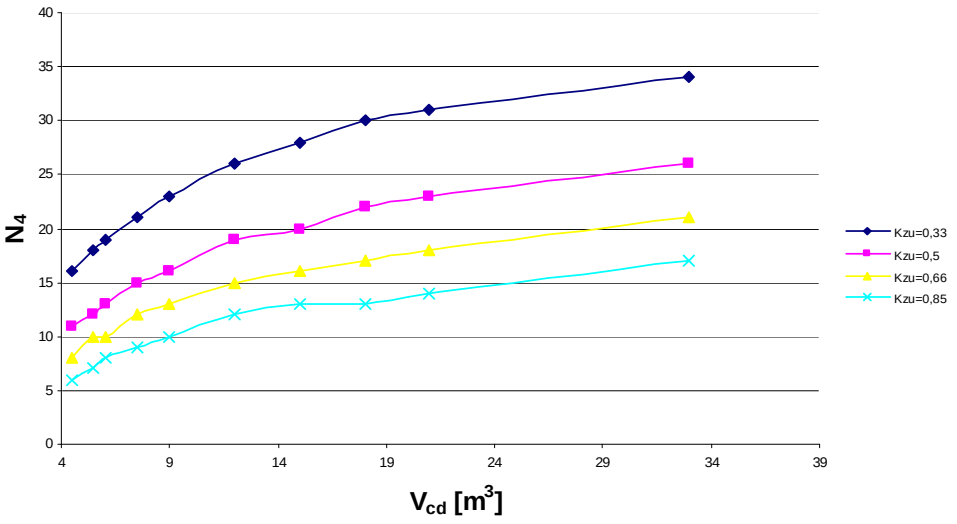


Fig. 2. The relation between the cisterns capacity and the number of refuelling cycles
Rys. 2. Wpływ pojemności cysterny na liczbę tankowań

The curves on Graph 2 show the relation between the number of refuelling cycles and the storage tank capacity. Equation (15) describes the mutual dependence:

$$N_4 = a + b * (\log cV_{cd} + d) \tag{15}$$

For a= 13.2647, b=4.705904, c=22.31486, d=-26.1808, the correlation coefficient is R^2 =98.89%, which means that the empirical curie fits well.

Summary

The article presents a method for the estimation the required number of tankers supplying aircrafts with aviation fuel, which depends on the type and number of the aircraft and the flight structure. This method is an attempt to use a mathematical model enabling the combination of theory and practice. This study is the results of two conditions. Firstly, there is no mathematical model, despite the fact that the problem is known and obvious. Secondly, the number of storage tanks is estimated by only human experience. A small defect of the proposed model might be the fact that that the calculations must be done in stages, grouped according to the flight duration (the same fuel consumption indexes)- (See the numerical example above). As a result of this method we obtained the minimum number of storage tanks of a particular capacity that are able to deliver the required amount of fuel to a certain place. This method could be

modified by introducing the excess factor (the number of storage tank will be bigger then). It will enhance the reliability of the fuel delivery system, but the costs will increase.

References

[1] Bobrowski D., Modele i metody matematyczne teorii niezawodności, WNT, Warszawa 1985.
[2] Fisz M., Rachunek prawdopodobieństwa i statystyka matematyczna, PWN, 1967.
[3] Gniedenko B.W., Bielajew J.K., Sołowiew A.D., Metody matematyczne w teorii niezawodności, WNT, Warszawa 1968.
[4] Ziółkowski J., Analiza systemu logistycznego bazy lotniczej w aspekcie gotowości, rozprawa doktorska, ITWL, Warszawa 2004.

Manuscript received by Editorial Board, August 31st, 2010

Metoda obliczania niezbędnej liczby cystern do zaopatrywania statków powietrznych w paliwo lotnicze

S t r e s z c z e n i e

W czasie wykonywania lotów przez statki powietrzne zasadniczymi kwestiami są gotowość bojowa oraz mający na nią wpływ system zaopatrywania. Zasadniczymi przedmiotami zaopatrzenia w czasie prowadzenia działań są środki bojowe, energia i paliwa lotnicze. Skuteczne zarządzanie przepływem wymaganych produktów a także niezawodność pojazdów i dysprozycyjność załogi w systemie logistycznym bazy lotniczej wpływa na jakość prowadzonych działań. Jakość ta w odniesieniu do cystern dystrybutorów może być mierzona m.in. niezawodnością dostaw i czynnikami ekonomicznymi (kosztami ich eksploatacji). Liczba cystern zaopatrująca statki powietrzne w paliwa lotnicze w czasie wykonywania lotów określana jest empirycznie. W artykule przedstawiono metodę wyznaczania niezbędnej (minimalnej) liczby cystern-dystrybutorów w zależności od rodzaju statków powietrznych, ich liczby, długości lotu oraz organizacji lotów.

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

STANISŁAW DUER*

Examination of the effectiveness of the operation process of a radar system

Key words

Servicing process, system modelling, expert system, artificial neural network, knowledge base, diagnostic information.

Słowa kluczowe

Proces obsługiwania, modelowanie systemów, systemy ekspertowe, sztuczne sieci neuronowe, bazy wiedzy, informacja diagnostyczna.

Summary

This paper presents a method to investigate the effectiveness of an operation process with a servicing expert system including an artificial neural network. A method of simulation testing with the use of computer technology is described. The theoretical bases are presented of the modelling of an operation process of objects in the form of the following models: mathematical (analytical), graphical and descriptional. For the tests, a model was developed of an organisation of a servicing system of those technical objects that require short shutdown times (aircraft, radiolocation systems, etc.). The requirements are presented and described for simulation tests, which is the development of a test plan, the preparation of data to describe the performance of an object, and the development of models for an operation process of a technical object, which express the investigated aspects of this process. The results are presented of the simulation tests of a repairable technical object.

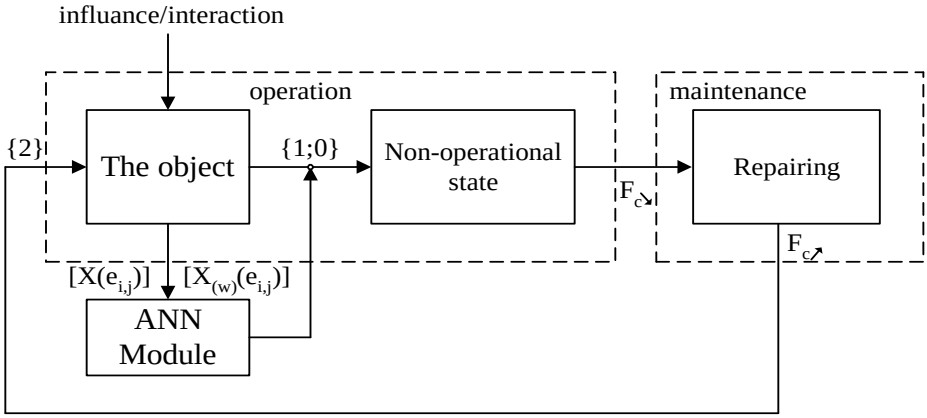
* Technical University of Koszalin, Department of Mechanics, Racławicka 15-17, 75-620 Koszalin, Poland, e-mail: stanislaw.duer@tu.koszalin.pl.

1. Introduction

Technical objects used in the process of operation may be found in specific states (operation or servicing) [1, 2, 10, 11, 12]. During operation, a technical object is exposed to various reactions, there are ageing changes, and the functional elements of the object wear. The changes of the physical properties of the object in relation to their nominal values result in the lowering value of the quality of functioning, thus the lowering value of the object's operational function during its operation. Then, the technical object passes to a shutdown state in the operation process (Fig. 1), which is determined by the states of non-operation or incomplete operation states.

In the shutdown state, the objects do not realise their foreseen operational functions. In order to counteract random changes of the object's quality of operation and to maximise it, the object's technical servicing is organised through planned and unplanned servicing.

Operation of technical (and other) objects is connected with the costs of operation. The user of a given technical object makes efforts to optimise the costs of operation and these issues are complex. In order to do so, one needs to establish the indices that characterise the operation process of a technical object, establish the performance of a given object (how to operate this object effectively, in which conditions it is to be operated, when servicing is to be organised and how to optimise this process, etc.), and also conduct tests of the operation process in the perspective of its effectiveness [1, 2, 3, 7, 8, 10, 11, 12].



where: $X(e_{i,j})$ – diagnostic signal in j^{th} element of ith set; $X_{(w)}(e_{i,j})$ – model signal for $X(e_{i,j})$ signal; F_C – function of the use of the object, $W(\varepsilon(e_{i,j}) = \{2, 1, 0\})$ – diagnostic information-value of state assessment logics for element „j" within „i" module of the object.

Fig. 1. Diagram of operation process for technical object utilising artificial neural network
Rys. 1. Schemat procesu eksploatacji wykorzystującego sztuczną sieć neuronową

2. Indices which characterize the operation process of a technical object

From the set presented in the literature [11, 12] of those indices which characterise the operation process of a technical object, the quantity which best reflects the operation process is availability factor K_g and availability function $K_g(t)$. The calculation process of availability function $K_g(t)$ is usually simplified when it is calculated for the boundary value at $t \to \infty$. This quantity has a close relation with the stationary characteristics of the process of damages and servicing. For this reason, availability factor K_g is the most suitable measure to express the effectiveness of the operation process, which combines at the same time the object's performance and economic properties.

Availability factor K_g of an object is the probability of an event that consists in the fact that the object is in working order after a sufficiently long period of operation ($t \to \infty$). Availability factor K_g determines the average share of the operation periods of a technical object in the total period of its operation, which is presented with the following dependence:

$$K_g = \lim_{t \to \infty} K_g(t) = \lim_{t \to \infty} K_{gsr}(t) \tag{1}$$

Where: K_{gsr} (t) – the average value of availability factor K_g.

The determination of availability factor K_g requires that the operation process of a given object be known exactly. While having the determined quantity that expresses the effectiveness of the operation process of an object of any class, one can determine the quality function of the operation process of the object.

The quality function of the object's operation process F_c is a dependence that characterises the object's operation process with respect to the effectiveness of the object's operation process in relation to the quantity of the input used during the total object's operation period, which is presented in the form of the following dependence:

$$F_c = \frac{K_g}{N_e} \tag{2}$$

The quality function of the object's operation process expresses the effects obtained in the process of the object's operation in the form of availability factor K_g in relation to operation input N_e borne in a given time (t) of the object's operation.

3. Modelling of the operation process of a repairable technical object

The operation process of a complex technical object is a stochastic process Z(t), whose elements (Z_i) belong to the subsets of the object's states {Z}: operation and servicing. While analysing any possible operation situations in which the object could be found after any number of passes, one can determine the object's states in the operation process, which form a set of the object's states Z(t). Each of the possible object's states in the operation process that can occur is determined in the diagnosing process of the object [1, 2, 4, 5, 6, 9]. The set of the object's states in the object's operation process {Z}, depending of the accepted valence of the valuation of states, is divided into two or three subsets {Z_1, Z_2, Z_3}, where: Z_1 – subset of operation states, Z_2 - subset of non-operation states, Z_3 - subset of incomplete operation states. The subset of the states of the operation of the operation process is a single-element process of the states of use Z_1 of the object in the operation process.

The state in which the object is used in the operation process [1, 6] is such a state of the operation process Z_1 during which the object realises the required functions in compliance with its use. If the object is no longer operated because it is not realising its required functions, then it should be repaired in the process of servicing. The states distinguished in the diagnosing process, when the object is not use, belong to the subset of the object's servicing states {Z_2}. This subset is a multi-element set, to which the following states belong: Z_2, Z_3, Z_4 where: Z_2 – planned servicing; Z_3 - unplanned servicing; Z_4 – ineffective use (the shutdown state). Depending of the need of research in the literature [3, 4, 7, 11], various methods are used to present the operation process of a technical object. Most frequently, the model of the operation of the object is represented in a graphical form. A graph of the process is the graphical form of the realisation of the operation process of the object. An analytical form is another manner to present the realisation of the operation process of the object.

3.1. The servicing technical object-the radar system type Straight Flush Radar Vehicle

In order to design the servicing system for an analogue class technical object [11], in this case it was an air-defence radar device, it was needed to determine the internal structure of serviced object and the set of prophylaxis activities for the non-operational elements. The radar system-Straight Flush Radar Vehicle presented in Fig. 2 is a part of a surface–to–air–missile system (SA-6 "GAINFUL"). The purpose of the radar system is to fight air targets (aircraft, helicopters, rockets, drone vehicles), as well as ground and water targets in the range of the missiles. The radar system, Straight Flush Radar Vehicle, detects (determines the azimuth, distance and height) and controls the air fight. The antiaircraft set is adapted to work regardless of the time of the

year and the day in temperatures from -40°C to +50°C, with a relative humidity of 90 per cent and the wind speed up to 20 m/s.

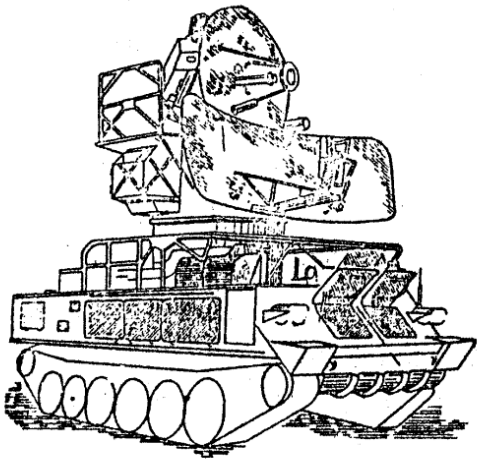
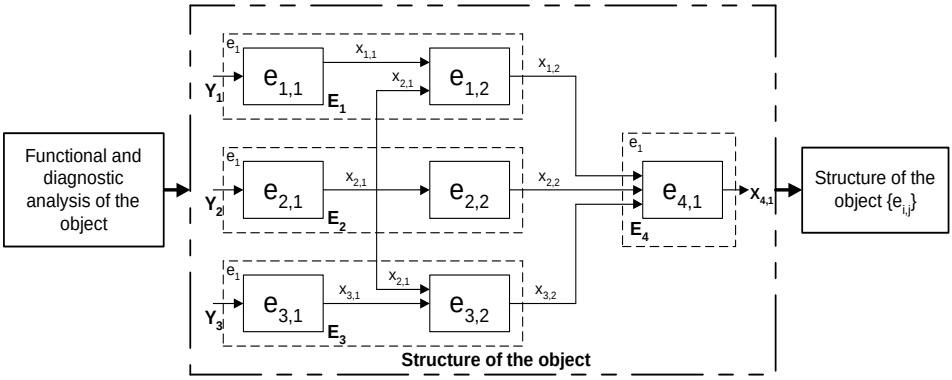


Fig. 2. The picture of an air-defence radar device of (SA-6) system
Rys. 2. Widok zestawu stacji radiolokacyjnej wykrywania i naprowadzania systemu (SA-6)

The radar system, Straight Flush Radar Vehicle, is characterised by a high resistance to climactic and natural factors. The (SA-6) system can be operated fully automatically as regards detection, identification, tracking, and raking of targets. The set is adapted to cooperation and coupling with four sources of external information. The radar can rake at the same time 1 target with 1 or 2 rockets fired within a span of 5 seconds from one or two launchers.

3.1. Modelling of the servicing object of a radar system, Straight Flush Radar Vehicle

The technical object used for tests in the present study is a reparable complex technical object of an analogue class. While preparing a diagnostic model of this class of an object, its internal structure was divided into four levels of the maintenance structure (Fig. 3): level one – object, level two – assemblies (in object), level three – subassemblies (in each assembly $\{E_i\}$), and level four – modules-basic elements (in each subassembly, of each assembly of the object). The first level of the maintenance structure of the object is constituted by the object itself. It is a set of functional assemblies $\{E_i\}$.



where: E_i - ijth functional assembly in the object,
e_i - j jth subassembly or functional element in a given assembly, $Y_{1,2,3}$ – input signals in the object.

Fig. 3. Functional and diagnostic model of the object
Rys. 3. Model funkcjonalno-diagnostyczny obiektu technicznego

Each functional subassembly of the object consists of basic elements, which are the smallest and indivisible functional element in the object. It was assumed in the paper that such an element is understood as a basic element in the object where there is an output (diagnostic) signal on its output. If object has been divided into *i* structural levels, and in each of them, there are *j* basic elements, then each of the object's structural levels constitutes a set of operating elements $\{e_{i,j}\}$, which was presented in the form of the following dependence:

$$\{O\} \Rightarrow (\{E_i\} \Rightarrow \{e_j\}) = \{e_{i,j}\} \tag{3}$$

Where: {O} – object's internal structure, $\Rightarrow$ – relation of result (division), E_i – ith functional assembly of the object, e_j – jth subassembly in ith assembly of the object, $\{e_{i,j}\}$ – set of basic elements in the object (structure of the object).

The division of the object's internal structure $\{e_{i,j}\}$ accepted in this paper defines explicitly the depth of penetration into this structure. The accepted division is considered to be sufficient if we distinguish the basic module-element in the structure of the object. One of the purposes of the functional-diagnostic analysis is the determination of the state of the object. The object's state is determined on the basis of an examination of the set of output (diagnostic) signals $\{X(e_{i,j})\}$(Table 1) [1-11]. The set of its functional elements $\{e_{i,j}\}$ determined during a diagnostic study of the object constitutes the basis for the list included in the table of a set of diagnostic signals (Table 1).

Table 1. Table of the input diagnostic signals of the object
Tablica 1. Tablica sygnałów diagnostycznych obiektu

Object	Level of object E_i	Vector of initial diagnostic signals $\{X(e_{i,j})\}$				
		$X(e_{1,1})$	...	$X(e_{i,j})$	...	$X(e_{i,J})$
O	E_1	$X(e_{1,1})$	...	$X(e_{1,j})$	...	$X(e_{1,J})$
	⋮	⋮	...	⋮	...	⋮
	E_i	$X(e_{i,1})$	...	$X(e_{i,j})$	...	$X(e_{i,J})$
	⋮	⋮	...	⋮	...	⋮
	E_I	$X(e_{I,1})$	...	$X(e_{I,j})$	...	$X(e_{I,J})$

Where: $X(e_{i,j})$ – diagnostic signal of jth element in i^{th} assembly.

3.2. Model of an expert servicing system with an artificial neural network of a radar system

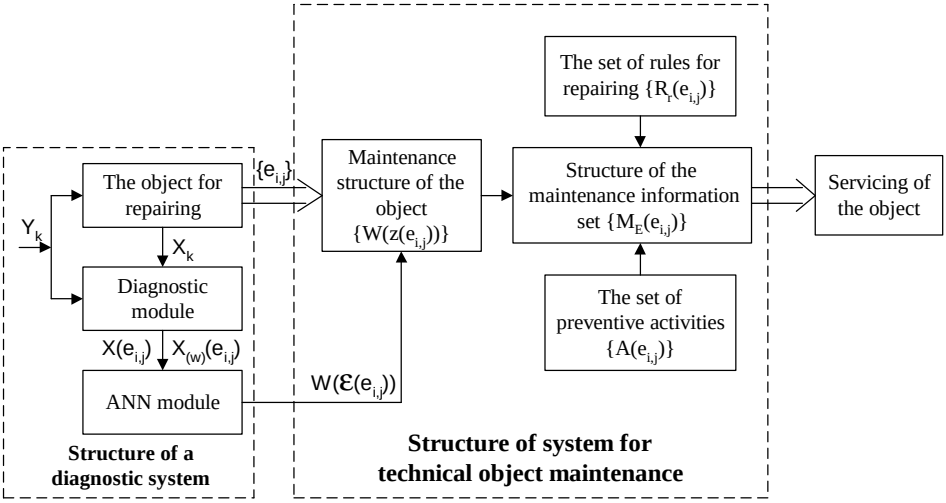
A particularly important element of the maintenance system is the knowledge base (Fig. 4). It can be defined as a specialised set of the object's maintenance information, which is determined by the following: the maintenance structure of the object $\{W_z(e_{i,j})\}$, the set of rules for maintenance (repairing) $\{R_r\}$, and the set of preventive activities $\{A(e_{i,j})\}$.

The quality of the use of an object can be measured with two quantities: the use function of the object $F_C(t)$ and F_C index of the use function of the object. The values of function $F_C(t)$ are determined by the divergence between the actual state of the object in the space of the use features (ω), and the state of the usability in the nominal space of usability features (M_E) (Fig. 1). The nominal space of usability features (M_E) is determined by elementary nominal vectors of the object's usability function $F_C(e_{i,j})$.

It is evident from the analysis (Fig. 4) that the process of the renovation of an object in an analytical approach consists in the transfer of the object's usability features from the level of the plane of the current use (ω) to the level of the plane of nominal usability features (M_E). The function that renovates the object in the servicing system is presented in the form of the following dependence:

$$M_E(e_{i,j}) = f\left(W\left(z(e_{i,j})\right); A_l(e_{i,j}); R_r(e_{i,j})\right) \tag{4}$$

Where: $\{W_z(e_{i,j})$ – the maintenance structure of the object $\}$, $\{R_r(e_{i,j})\}$ – the set of rules for maintenance (repairing), $\{A(e_{i,j})\}$ – the set of preventive activities, $\{M_E(e_{i,j})\}$ – the maintenance system produces a set of maintenance information.



Where: $X(e_{i,j})$ – diagnostic signal in j^{th} element of ith set; $X_{(w)}(e_{i,j})$ – model signal for $X(e_{i,j})$ signal; F_C – function of the use of the object.

Fig. 4. Diagram of a servicing system with an artificial neural network
Rys. 4. Schemat procesu obsługiwania obiektu technicznego wykorzystującego sztuczną sieć neuronową

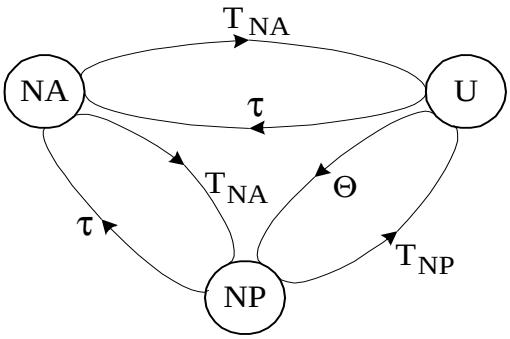
As a result, the maintenance system produces a set of maintenance information $\{M_E(e_{i,j})\}$, which will be used for the organisation of the object's rational (optimised) maintenance system.

A set of the servicing information of the object, which constitutes the basis in the process of designing of the structure of the servicing system, is presented in the form of the following dependence:

$$M_E(e_{i,j}) = \{M(e_{i,j}), W(z(e_{i,j}))A_l, R_r\} \tag{5}$$

3.3. Modelling of the operation process of a technical object

Every operation process of any object consists of a subset of the operation state and a subset of the servicing state (Fig. 5). The description of passes between the distinguished states of an object's operation process tested, and relations between them constitutes a difficult task and requires that the real operation process is known exactly.



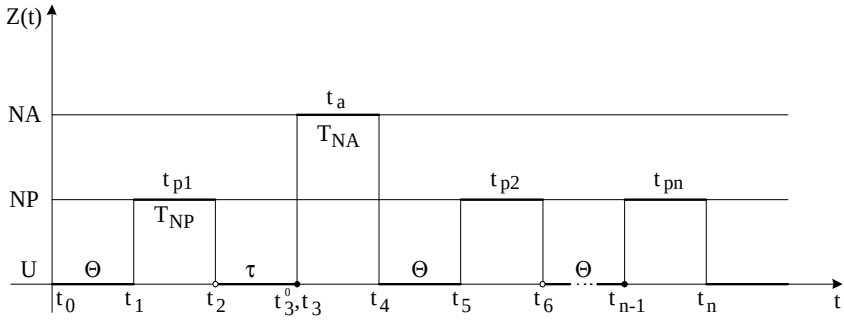
Where: T_{NA} – duration of unplanned repair; T_{NP} – duration of planned repair; Θ – periodicity of the object's planned repair, τ – intensity of damages.

Fig. 5. Graph of a model of an object's operation process with a servicing system which uses an artificial neural network which processes diagnostic information expressed in the values of the three–values {2, 1, 0} of the states
Rys. 5. Graf modelu procesu eksploatacji obiektu z systemem obsługiwania wykorzystującym sztuczną sieć neuronową wypracowującą informację diagnostyczną wyrazaną w trójwartościowej ocenie stanów

Only when it is known, a model of the operation process of the object can be developed while one of the ways are used for this purpose to present a given structure of the object's operation process. In each operation process of the object described, the most difficult tasks are as follows: the development of the model of the structure of the object's operation, as well as establishing and describing a possible subset of the state of the technical object's operation.

While having a constructed servicing system (Fig. 5) and knowing the relations between operation states, one can begin to develop a model of the object's operation process under investigation. Its practical realisation is presented in Fig. 5. In order to understand the idea of the operation process model developed, one needs to analyse the mechanism of the change of the states (the way in which passes between the states occur). A practical manner of the organisation of a pass between the distinguished operation states of the object in its developed model is presented in Fig. 6.

An analysis of the ways of the change of the states of the object operated in compliance with the model described demonstrates that this object can be found in of the following states:
– Z_1 – operation U;
– Z_2 – scheduled servicing (prophylactic repair);
– Z_3 – unplanned repair NA (emergency).



Where: • – moments of object's damages; ° – moments of object's repairs.

Fig. 6. Example of the realisation of the object's operation process model with a servicing system using an artificial neural network which develops diagnostic information expressed in the values of the three-values {2, 1, 0} of the states
Rys. 6. Przykład realizacji modelu procesu eksploatacji obiektu z systemem obsługiwania wykorzystującym sztuczną sieć neuronową wypracowującą informację diagnostyczną wyrażaną w trójwartościowej ocenie stanów {2, 1, 0}

From among the operation indices presented above, the availability factor was considered to be the most dependant, from the changes of the operation process quality, and which is described with the following dependence:

$$K_g(\Theta) = \frac{\int_0^{\Theta} R(t)dt}{\int_0^{\Theta} R(t)dt + T_{NP} \cdot F(\Theta)} \tag{6}$$

Where: T_{NP} – mean duration of the object's scheduled repair; $F(\Theta) = P\{ł < \Theta\}$ – value of the distribution function of time until damage in moment $t = \Theta$; $R(\Theta) = P\{ł \geq \Theta\} = 1 - F(\Theta)$ – value of survival function in moment $t = \Theta$.

It is evident from an analysis of dependence (6) that availability factor K_g can be determined when value Θ is known which determines the periodicity of the realisation of the object's scheduled repair. If $\lambda(\Theta^*) = \Theta^*$ is a function constant in time, then dependence (6) accepts the following form:

$$K_g = \frac{T}{1 + (T_a - T_p) \cdot \Theta^*} \tag{7}$$

Where: T – operation time, T_a – emergency repair time, T_p – prophylaxis time, Θ^* – optimal periodicity of prophylactic servicing.

Considering the need to relate future results of the research in a comparative analysis of the models of operation processes, a model must be developed of an operation process of a real object with a traditional servicing system.

4. Simulation test of the effectiveness of a radar system's operation process

An investigation of the operation process of an object, and all the more a simulation investigation, requires input data which characterises the real operation process of a selected class of an object and its simulation models. The results of simulation tests obtained for the accepted duration of the investigation T_o are presented in (Figs. 7 and 8).

An investigation of the real operation process of an object constitutes the basis for obtaining data for a simulation investigation of the models of processes. The following quantities constitute the required input data for tests:

- Duration of the use of the object T (time when the object is in the operation state);
- Duration of the removal of the object's non-operation state T_a;
- Duration of prophylactic repair T_p;
- The period of forecast (optimal) prophylaxis θ^*; and,
- The period of planned prophylaxis θ.

The operating quantities used for the purpose of a simulation examination of the object's models of operation processes were determined on the basis of observations and an examination of the technical and operating documentation of the device tested of the same class [11]. The input data accepted for the investigations are as follows:

- The average time of the removal of the non-operation state (T_a) is determined from (T/T_a) relation for given values included in set {B}, where T/T_a = {B}; for the examinations, values {B = 1, 2, 4, 8} were accepted; and,
- The average time of a prophylactic repair (T_p) is determined from (T/T_p) relation for simulation examinations; a constant value (5) was accepted, where: (T/T_p) = 5; the average period of the planning of a prophylactic repair was determined in the diagnostic experiment in the way of forecasting of the time of further preventive treatment, and this time amounts to θ^* = 115 [h - time units] [11].

The source of the data mentioned above may be an observation of real operation processes as well as a suitably prepared and conducted simulation experiment.

The simulation investigations were carried out for constant inputs for prophylaxis in the whole period of operation (a simulation time of the use of the

object: T_o). As an evaluation of a real operation process is too time consuming, three operation process models were developed for the purpose of investigations:

1. Model A: an object's operation process using information in trivalent logic (a model with an expert system, with artificial neural network),
2. Model B: an object's operation process using information in divalent logic,
3. Model C: an operation process of an object organised in a classical manner, i.e. without an investigation of the state.

In the investigations, the following were evaluated: the indices which characterise the operation process of technical objects, availability factor K_g and the quality function of the operation of the object process F_c. The determination of the value of availability factor K_g and the quality function of the object's operation process requires that the properties of the operation process of the technical object used are known exactly: the characteristics of the damage and servicing processes (T – time of use and T_a – time of servicing).

There are the following conclusions from the analysis (Figs. 7 and 8) of dependencies (K_g) and (F_C) expressed in the function of test duration (T_o) – conventional units of the test duration:

1. The results of the investigation of quantities (K_g) (Fig. 7) for the model (A and B) are decreasing, and the largest value of (K_g) is for model A, and it is (K_{gA} = 0.7508). For the remaining models, this quantity is as follows: (K_{gB} = 0.4931 and (K_{gC} = 0.2332).
2. The quantity investigated (K_g) has also a practical aspect: it refers to the effectiveness of the organisation of a given servicing system. We may find from the definition of this quantity: dependence (1) of quantity (K_g) [1 -11] that this factor determines the availability of the object to be used as it is intended for use in the period of use (T). Therefore, the organisation system of prophylaxis is the most effective presented in model (A), as value (K_{gA} = 0.7508) is the maximum value.
3. The tested function of the operation process quality (F_C) in models (A, B, and C), that is graphically represented in (Fig. 8) is also expressed in the accepted time of the test (T_o). It is evident from the definition of quantity (F_C): dependence (2) that the input borne in the object's servicing system organisation in a given model has a large impact on this quantity.
4. The quantity of inputs for models (A, B, and C) and a constant expressed with a linear function were accepted for the tests. It is evident from the analysis (Fig. 8) that the best quality of operation process (F_C) is for model (B), where: (F_{CB} = 0.2239), while for model (A) it is (F_{CA} = 0.1979). It can be supposed that the organisation of the servicing system in model (A) is more cost consuming, and this is a more technologically advanced servicing system.

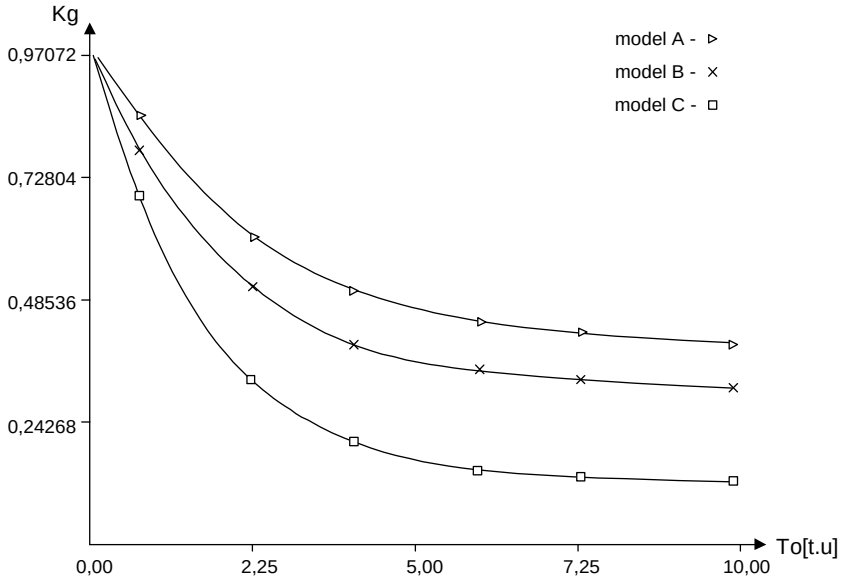


Fig. 7. Availability function $K_g(t)$ of operation process models tested (for linear functions of expenditures) for T/Ta = 2 and $N/T_o = 1$

Rys. 7. Funkcja gotowości $K_g(t)$ badanych modeli procesu eksploatacji (dla liniowej funkcji nakładów dla T/Ta = 2 and $N/T_o = 1$

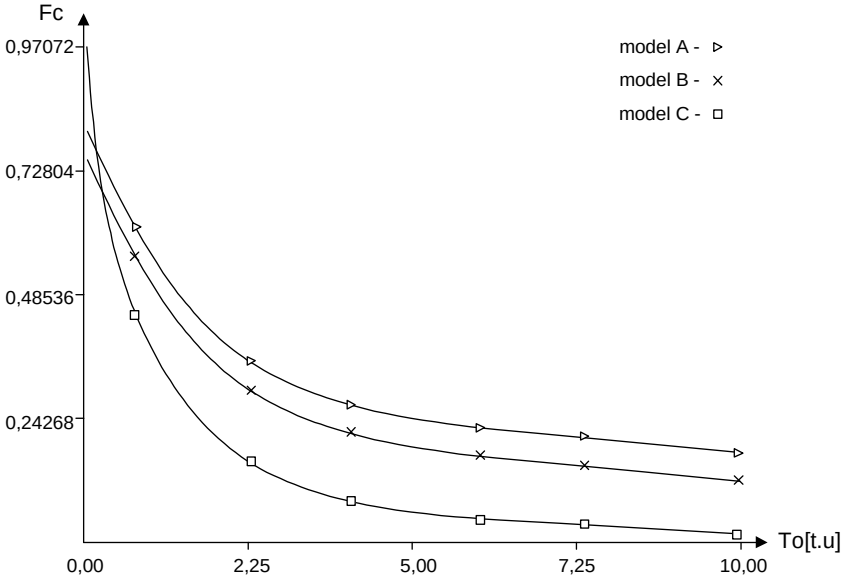


Fig. 8. Quality function of object's operation process F_c of operation process models tested (for linear functions of expenditures) for T/Ta = 2 and $N/T_o = 1$

Rys. 8. Funkcja jakości procesu eksploatacji obiektu F_c badanych modeli procesu eksploatacji (dla liniowej funkcji nakładów dla T/Ta = 2 and $N/T_o = 1$

The results of the investigations conducted serve to confirm the justifiability of the accepted concept of the organisation (designing) of the servicing system of repairable technical objects with the required short shutdown time.

In the organisation process of every servicing system, the costs constitute the factor that determines this process. The results obtained in the research indicate that an optimisation of the costs of the prophylaxis of technical objects can be achieved in two ways. The first one of them indicates that only those elements of an object should be renovated which require that these are elements in the state of an incomplete operation and in the non-operation state (therefore, there is a problem of the designing of the structure of the servicing system). According to the second manner of decreasing the costs of regeneration, the object's elements need to be regenerated in a selective manner by performing not all of the regeneration activities but only prophylactic activities as selected by an expert (an expert knowledge base), which are suitable for the state in which a given element is of an object of a given class of the servicing element.

The issues of the modelling of operation processes of technical objects are complex. They are related to many fields of science, such as mathematics, theory of operation, reliability theory, technical diagnostics, theory of systems modelling, information technology, etc. Each of these fields is intensively being developed at present.

5. Conclusions

This study covers a method of a simulation testing of the models of the organisation of servicing of technical objects with the required short time of their shutdown (aeroplanes, radiolocation systems, etc.). The issues presented in the study concerning the testing of the performance of technical objects during operation present a difficult task. This difficulty is the result of the length of the time for testing in the real process of the object's operation (this is the time of the "life" of the object), which is practically expressed in decades of years. The only reasonable approach for this type of research is simulation testing. However, this requires the real operation process of a given technical object to be known, on the basis of which the following needs to be prepared: the model of the operation process of a given technical object, plan of testing, input data for the test to describe the performance of a given object, such as the object's use time T - the time when the object is in the operation state, T_a – time of the removal of the object's non-operation state, T_p – time required for the performance of a prophylactic repair, the specific nature of the prophylaxis of this class of objects.

The models of the servicing system organisation constitute the basis of a simulation testing of the operation process of a technical object. The study presents the idea of one example model of the operation process with the

servicing system tested, which makes use of diagnoses in a trivalent logic. The model presented in the article of the organisation of a servicing system which makes use of diagnoses in a trivalent logic is an innovative and interesting solution.

References

[1] Duer S., Duer R.: Computer diagnostic system with expert knowledge for state's control of technical objects. "SYSTEM MODELLING CONTROL – SMC'2005". Proceedings of the 11th International Conference on "System Modelling Control", Zakopane, Poland. Wyd. EXIT, Warszawa 2005, pp. 69-77.

[2] Duer S.: The concept of assistant system for analogue class technical object servicing. Sixth International Conference On Unconventional Elektromechanical And Electrical System UEES′04. Alushta, The Crimea, Ukraine, 2004, pp. 687–690.

[3] Duer S.: System ekspertowy wykorzystujący trójwartościową informację diagnostyczną wspomagający obsługiwanie złożonego obiektu technicznego. Zagadnienia Eksploatacji Maszyn Z. 4(152) VOL. 42, 2007, pp. 195-208.

[4] Duer S.: An algorithm for the diagnosis of reparable technical objects utilizing artificial neural Network. Scientific Problems Of Machines Operation And Maintenance, Committee Of Machine Engineering Polish Academy Of Sciences. Vol. 43, No. 1(53) 2008, pp. 101––113.

[5] Duer S., Duer R., Duer P., Płocha I.: Measurement system for the diagnosis of analogue technical objects with the use of artificial neural networks. Academic Journals, Poznan University of Technology, s. Electrical Engineering, Published by Poznan University of Technology, No. (59), Poznan 2009, pp. 61-72.

[6] Duer S.: Determination of a diagnostic information of a reparable technical object on the basis of a functional and diagnostic analysis on example of a car engine. Scientific Problems Of Machines Operation And Maintenance, Committee Of Machine Engineering Polish Academy Of Sciences. Vol. 43, No. 4(156) 2008, pp. 85-94.

[7] Duer S.: Classification of the elements of the object's maintenance structure of a reparable technical object. Scientific Problems Of Machines Operation And Maintenance, Committee Of Machine Engineering Polish Academy Of Sciences. 2009 (in publishing).

[8] Duer S.: System ekspertowy, ze sztuczną siecią neuronową obsługujący zestaw stacji radiolokacyjnej. VII Krajowa Konferencja Inżynierii Wiedzy i Systemów Ekspertowych, Politechnika Wrocławska, 23 – 25 Wrocław, 2009, (in publishing).

[9] Duer S.: Artificial Neural Network-based technique for operation process control of a technical object. Defence Science Journal, Vol. 59, No. 3, May 2009, pp. 305-313, DESIDOC.

[10] Duer S.: Creation of the servicing information to support the maintenance of a technical object with the use of three-value logic diagnostic information. Scientific Problems Of Machines Operation And Maintenance, Committee Of Machine Engineering Polish Academy Of Sciences. 2009 (in publishing).

[11] Łosiński M., Duer S., Jaśkiewicz A., Drabarek J., Madej W.: Analiza efektywności procesu eksploatacji sprzętu przeciwlotniczego. WSOWOPL nr wewn. 1035/94. Koszalin 1994.

[12] Nakagawa T.: Maintenance Theory of Reliability. Springer – Verlag London Limited 2005, p. 264.

Manuscript received by Editorial Board, July 7th, 2010

Badanie efektywności procesu eksploatacji zestawu stacji radiolokacyjnej

S t r e s z c z e n i e

W pracy zaprezentowano metodę badania efektywności procesu eksploatacji, w którym występuje obsługowy system ekspertowy, ze sztuczną siecią neuronową. Opisano sposób badania symulacyjnego z wykorzystaniem techniki komputerowej. Przedstawiono podstawy teoretyczne modelowania procesu eksploatacji obiektów w postaci modelu: matematycznego (analitycznego), graficznego i opisowego. Na potrzeby badań opracowano model organizacji systemu obsługiwania obiektów technicznych wymagających krótkiego czasu ich przestojów (samoloty, systemy radiolokacyjne itp.). Przedstawiono i opisano wymagania do badania symulacyjnego, którymi są: opracowanie planu badania, przygotowania danych opisujących właściwości użytkowe obiektu oraz opracowanie modeli procesu eksploatacji obiektu technicznego wyrażających badane aspekty tego procesu. Przedstawiono wyniki badania symulacyjnego naprawialnego obiektu technicznego.

SCIENTIFIC PROBLEMS
OF MACHINES OPERATION
AND MAINTENANCE
3 (163) 2010

ANDRZEJ KATUNIN*

B-spline wavelet packets and their application in the multiresolution non-stationary signal processing

Key words

Wavelet packets, multi-resolution signal processing, faults identification.

Słowa kluczowe

Pakiety falkowe, wielorozdzielcza analiza sygnałów, identyfikacja uszkodzeń.

Summary

Increasing requirements for the technical condition of machines induce the development of novel diagnostic methods for possible fault detection and identification in an early phase. Most of these methods are based on the processing of vibration signals. The classical methods often do not give full information about the actual condition of a machine. Therefore, it is necessary developing the appropriate diagnostic methods. Some of the promising signal processing methods are the group based on the Wavelet Transform (WT), which give a possibility for the effective diagnosing of non-stationary vibration signals in the time-scale domain. The generalisation of WT, the Wavelet Packet Transform (WPT), allows the extraction of additional useful diagnostic features from the signal. However, the effectiveness of diagnostics in the case of wavelet-based methods is determined by the selection of an appropriate wavelet function. In the present study, the author introduces new wavelet packets based on B-spline wavelets. A comparative analysis of their effectiveness was performed on non-stationary synthetic signals. The B-spline wavelet packets were applied for rolling bearing condition evaluation.

* Department of Fundamentals of Machinery Design, Faculty of Mechanical Engineering, Silesian University of Technology, Konarskiego 18A, 44-100 Gliwice, Poland, e-mail: andrzej.katunin@polsl.pl

1. Introduction

The progress in modern machine construction implies the development of diagnosing systems, which use increasingly complicated algorithms of the measured signals processing. Increasing requirements, according to the durability of machines and their low-cost maintenance, demand much improvement in signal processing methods with the following criteria: The method must give the most complete information about the technical condition of a machine, the method must detect and identify faults possibly clearly and it must allow faults detection and identification in possible early stage.

As it is well known, the most informative vibration signal in rotor machinery could be registered during machine run-up or run-out. Usually, these signals are non-stationary and the classical methods like Discrete Cosine Transform (DCT) or Fast Fourier Transform (FFT) are not efficient. The methods for non-stationary signal analysis are the Short-Time Fourier Transform (STFT) and order tracking analysis [1]. However, in the case of STFT, the basic function is the sine function, and it allows for only constant-resolution analysis. An additional and widely used method of signal processing is WT, which gives the possibility of multi-resolution signal analysis and the basic function is not limited to the one type.

Previous research indicates that the Discrete Wavelet Transform (DWT) allows the extraction of more diagnostic features from the vibration signal. The comparative analysis of FFT, STFT, and DWT effectiveness was performed for the structural damage detection of polymeric composites [2]. The most crucial factor that influences the effectiveness of WT is the appropriate wavelet selection. In the case of DWT, the selected wavelet must fulfil some additional criteria, e.g. compact support or orthogonality; thus, not every available wavelet could be applied for the analysis. The most popular wavelets used in the technical diagnostics for signal processing using DWT are the Daubechie's wavelets [3], biorthogonal wavelets, symlets, and coiflets [4,5].

Katunin's research shows the higher effectiveness of B-spline wavelets in comparison with the above-mentioned ones first applied in the diagnostic signal processing problems in [6]. In [7], the comparative study for some orthogonal wavelets was carried out using the degree of scalogram density (DSD) parameter. The analysis was carried out for the most typical signal components: the harmonic, the harmonic with variable frequency, and the pulse component. In all cases, the B-spline wavelets show excellent results. Furthermore, these wavelets were applied for the detection and localisation of single [7] and multiple [8] cracks in composite beams.

An additional wavelet-based method used in the diagnostic signal processing is WPT. This method, proposed by Coifman *et al*. in [9], is a kind of generalisation of DWT, which provides full multilevel decomposition of the signal. The graphical interpretation of DWT and WPT is presented in Fig.1.

WPT found many applications in mechanical and technical diagnostics problems. Glabisz, in his monograph [11], showed the application of WPT in signal processing problems, nonlinear systems identification, and the solution of differential equations. In diagnostic problems, WPT is often used for damage identification. The entropy-based method for rolling bearing condition diagnostics was developed by Wysogląd [3,12,13]. The rolling bearing condition could be analysed using other methods based on WPT, e.g. weak signature detection method [14,15]. WPT is also used in gearbox diagnostics [16], structural damage detection in multi-layered composites [17], medical applications [18] and others.

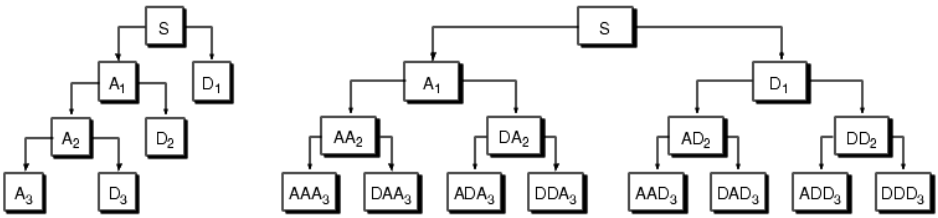


Fig. 1. The principle of multilevel decomposition using Discrete Wavelet Transform and Wavelet Packet Transform [10]
Rys. 1. Zasada wielopoziomowej dekompozycji z zastosowaniem Dyskretnej Transformaty Falkowej i Pakietowej Transformaty Falkowej [10]

Considering the excellent results for B-spline wavelets, the author decided to construct new wavelet packets based on these wavelets. Furthermore, the general order B-spline wavelet packets mathematical formulation will be proposed. Then, their effectiveness will be analysed in application for synthetic non-stationary signals based on Inverse Degree of Scalogram Density (IDSD) parameter. Finally, the B-spline wavelet packets will be applied for fault identification in rolling bearings, and the obtained results will be compared with the results obtained using other wavelet packet bases.

2. B-spline wavelets and wavelet packets

2.1. General order B-spline wavelets

The B-spline wavelets were introduced by Chui in [19]. The scaling function φ_m of the B-spline wavelet is defined by the two-scale relation:

$$\varphi_m(x) = \sum_{k=0}^{m} p_k \varphi_m(2x-k),\ k \in Z \tag{1}$$

Where

$$p_k = 2^{1-m}\binom{m}{k},\ 0 \le k < m \tag{2}$$

is the two-scale sequence.

The wavelet function ψ_m is also presented by the two-scale relation in the form of

$$\psi_m(x) = \sum_{k=0}^{3m-2} q_k \varphi_m(2x-k),\ k \in Z \tag{3}$$

Where

$$q_k = (-1)^k 2^{1-m} \sum_{l=0}^{m} \binom{m}{l} \varphi_{2m}(k-l+1),\ 0 \le k < m \tag{4}$$

The decomposition using B-spline wavelets could be carried out with the following relation:

$$\varphi_m(2x-l) = \sum_k \left(a_{l-2k}\varphi(x-k) + b_{l-2k}\psi(x-k)\right),\ l \in Z \tag{5}$$

Where

$$a_k = \frac{(-1)^{k+1}}{2} \sum_l q_{-k+2m-2l+1} c_{l,2m} \tag{6}$$

$$b_k = -\frac{(-1)^{k+1}}{2} \sum_l p_{-k+2m-2l+1} c_{l,2m} \tag{7}$$

are the decomposition sequences. The coefficient sequence $c_{k,m}$ could not be determined analytically for $m < 5$ [7]; thus, it could be calculated from the bi-infinite system of equations [19]:

$$\sum_{k=-\infty}^{\infty} c_{k,m} \varphi_m\left(\frac{m}{2} + j - k\right) = \delta_{j,0}\ ,\ j \in Z \tag{8}$$

The approximate solution of (8) and the method of the determination of decomposition coefficients $c_{k,m}$ could be found in [19]. The explicit form of B-spline wavelets with order of 2 – 4 could be found in [6] and with order of 5 – 7 are presented in [7]. Note that the first order B-spline wavelet is the Haar wavelet.

2.2. Construction of the general order B-spline wavelet packets

The Discrete Wavelet Packet Transform (DWPT) is a generalisation of DWT which yields time-scale decomposition by decomposing not only the approximate coefficients (as in DWT) but also the details coefficients (see the scheme in Fig. 1). It could improve the localisation of high frequency components. Following the dependencies presented by Coifman *et al.* [9] and considering two-scale relations for scaling (1) and wavelet (3) functions, the general order B-spline wavelet packets could be constructed by induction as a set of functions represented by the following recursive equations:

$$\begin{aligned} w_{2n,m}(x) &= \sum_k p_k w_{n,m}(2x-k),\ k \in Z \\ w_{2n+1,m}(x) &= \sum_k q_k w_{n,m}(2x-k),\ k \in Z \end{aligned} \tag{9}$$

Note that $w_{0,m} = \varphi_m$ and $w_{1,m} = \psi_m$. The B-spline wavelet packets have some unique properties, and some of them are related to B-spline wavelet properties presented, e.g. in [20]. The main properties are mentioned below:

- All of the B-spline wavelet packets are of compact support. From (9) it could be concluded that B-spline wavelet packets are compactly supported on the finite intervals. For the general order B-spline wavelet packets considering (9), we obtain:

$$\begin{aligned} \operatorname{supp} w_{0,m} &= [0, \operatorname{supp} \varphi_m] \\ \operatorname{supp} w_{n,m} &= [0, \operatorname{supp} \psi_m] \end{aligned},\ n, m \in Z,\ n = 1,2,... \tag{10}$$

- B-spline wavelet packets are symmetric for even m and any n. In the case of odd m, they are symmetric for even n and anti-symmetric for odd n, namely:

$$\begin{cases} w_{n,m}(x) = w_{n,m}(2m-x-1), & \text{for even } m \\ \begin{cases} w_{n,m}(x) = w_{n,m}(2m-x-1), & \text{for odd } m \text{ and even } n \\ w_{n,m}(x) = -w_{n,m}(2m-x-1), & \text{for odd } m \text{ and odd } n \end{cases} \end{cases} \tag{11}$$

- B-spline wavelet packets are semi-orthogonal, that is, they are orthogonal during scaling and not orthogonal during translation:

$$\left\langle w_{n,m}(2^i x - k_1), w_{n,m}(2^j x - k_2) \right\rangle = 0,\ i \neq j,\ k_1 \neq k_2 \tag{12}$$

$$\left\langle w_{n,m}\left(2^{j}x-k_1\right), w_{n,m}\left(2^{j}x-k_2\right)\right\rangle \neq 0\,,\ k_1 \neq k_2\,,\ m \neq 1\,.$$

Considering (5) and (9), the multilevel decomposition relation for discussed wavelet packets could be presented as follows:

$$w_{n,m}(2x-l)=\sum_{k}\left(a_{l-2k}\,w_{2n,m}(x-k)+b_{l-2k}\,w_{2n+1,m}(x-k)\right) \qquad l\in Z\,, \tag{13}$$

Where a_k and b_k are the decomposition sequences given by (6) and (7), respectively.

The B-spline wavelet packets constructed from the B-spline wavelets of order 3 (quadratic) and of order 7 (sextet) are presented in Fig. 2 and Fig. 3, respectively. In both cases, the parameter n is in the range of 0 - 7. Analysing (10), one could convince that the support of a given wavelet packet of order m equals the support of the appropriate wavelet function of order m. However, the length of the parts with highest disturbances of a given function is almost stable.

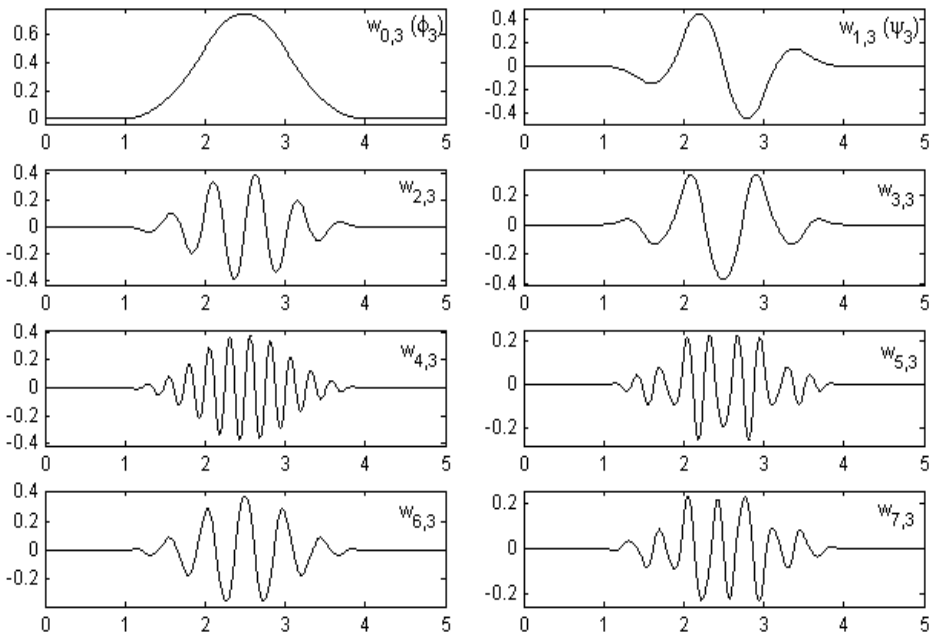


Fig. 2. Wavelet packets based on quadratic B-spline wavelet for n = 0 - 7
Rys. 2. Pakiety falkowe na podstawie falki B-splajnowej rzędu 3 dla n = 0 ÷ 7

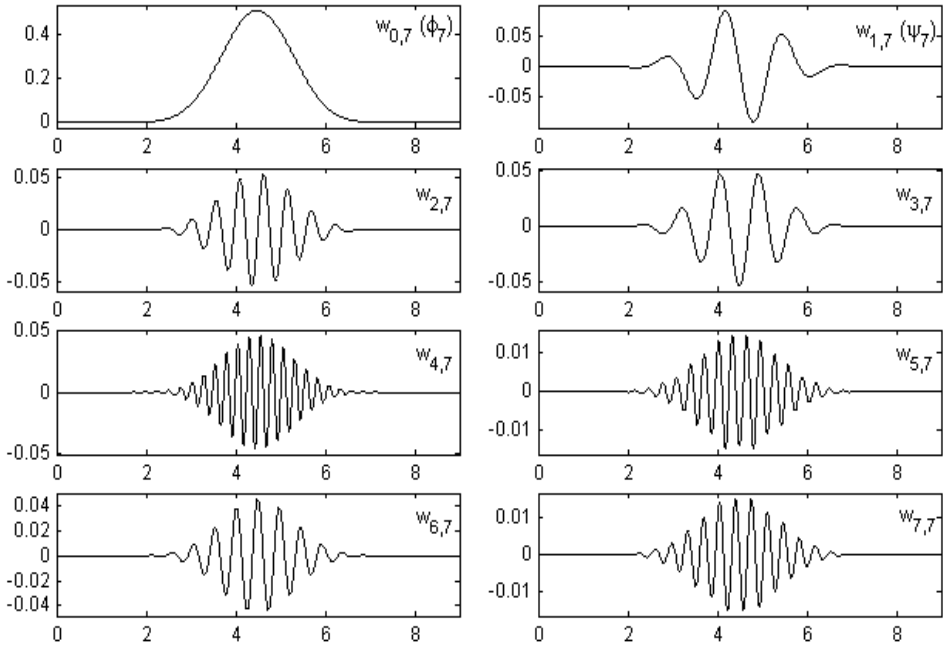


Fig. 3. Wavelet packets based on sextic B-spline wavelet for n = 0 - 7
Rys. 3. Pakiety falkowe na podstawie falki B-splajnowej rzędu 7 dla n = 0 ÷ 7

3. Effectiveness of the evaluation based on synthetic signals

For the evaluation of the wavelet packet effectiveness in the signal processing problems, the analysis based on synthetic signals using IDSD was performed. The DSD parameter used in [1, 6, 7] for the evaluation of the approximation effectiveness of wavelets is a scalar value, which is based on the determination of the ratio between the number of the wavelet coefficients higher than some non-zero threshold and the number of all wavelet coefficients in the scalogram. The IDSD parameter gives the ratio between the number of the wavelet coefficients higher than the threshold and the number of all wavelet coefficients. It makes possible the analysis of disturbances inputted by the given wavelet packets.

As previously mentioned, the best application of WPT is the analysis of non-stationary signals. In this study, three types of such signals were analysed: the linear chirp, the quadratic chirp, and the logarithmic chirp. The signals were synthetically generated with the length of 1 second and sampling rate of 1/512 s with the frequency range of 10 - 250 Hz. Then, the WPT algorithm was applied with the 4th decomposition level. Obtained coefficients were normalised into the closed set [0,1]. An exemplary comparison between scalograms with the Daubechies 3 (*db3*) wavelet packet base (1st column) and the B-spline 3 (*bsp3*)

wavelet packet base (2nd column) is presented in Fig.4 for the linear (1st row), quadratic (2nd row) and logarithmic (3rd row) chirp. On the vertical axes, the wavelet packet decomposition tree numbers were shown. The order of priority of the wavelet packets was set due to the frequency order.

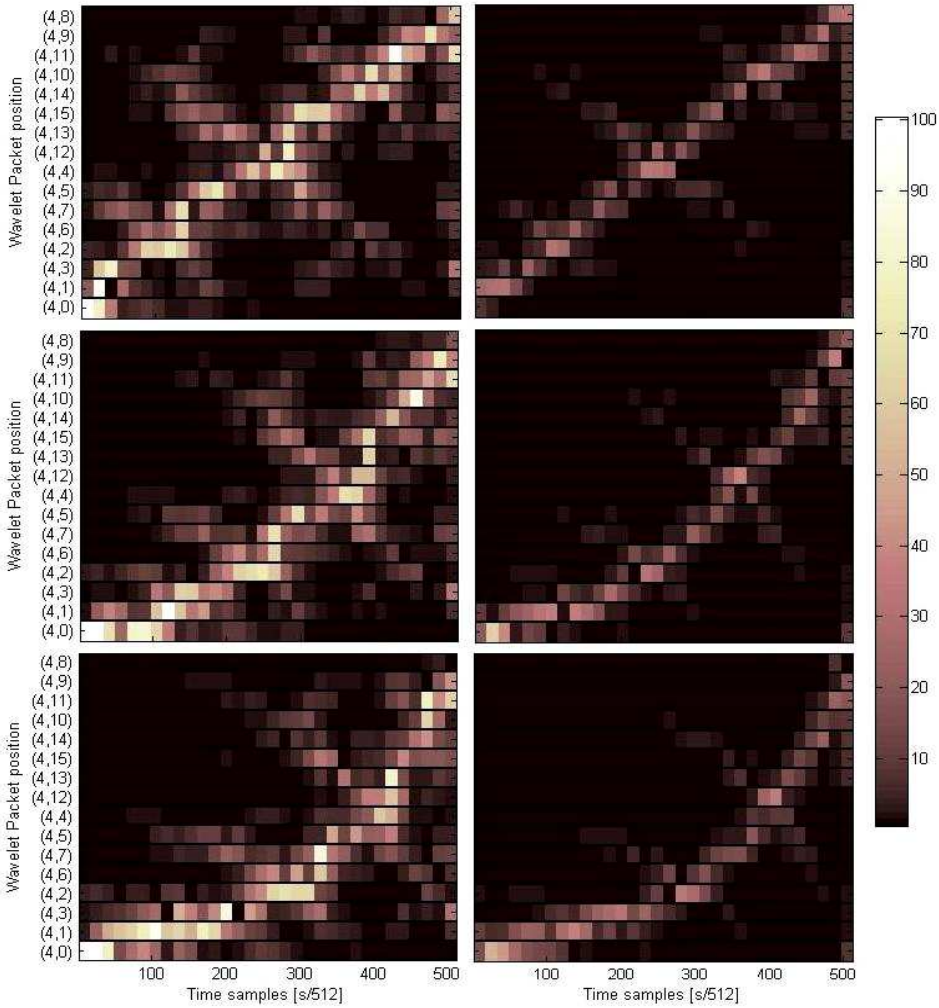


Fig. 4. Scalograms after WPT using Daubechies 3 and B-spline 3 bases
Rys. 4. Skalogramy po pakietowej transformacji falkowej z użyciem baz Daubechies 3 i B-splajn 3

Analysing Fig. 4 it can be noticed that the quadratic B-spline wavelet packet base gives lower disturbances in comparison with *db3* wavelet packet base for all of the investigated signal types. For determining the measure of the effectiveness of the wavelet packets, the IDSD was calculated for Daubechie's

wavelets (*db*), B-spline wavelets (*bsp*), symlets (*sym*) and coiflets (*coif*) with order from 1 to 5. Results of the calculation are presented in Table 1. The effectiveness of given wavelet packets is better when IDSD tends to the minimum.

Table 1. IDSD values for several wavelet packet bases
Tabela 1. Wartości odwrotnego stopnia zagęszczenia skalogramu dla różnych baz pakietów falkowych

	db1	**db2**	**db3**	**db4**	**db5**
Linear chirp	0.4414	0.3695	0.3090	0.2401	0.2656
Quadratic chirp	0.4531	0.3217	0.2743	0.2237	0.2344
Logarithmic chirp	0.4238	0.3419	0.2448	0.2237	0.2297
	bsp1	**bsp2**	**bsp3**	**bsp4**	**bsp5**
Linear chirp	0.4414	0.2554	0.1528	0.1014	0.0888
Quadratic chirp	0.4531	0.2339	0.1476	0.1047	0.0938
Logarithmic chirp	0.4238	0.2125	0.1389	0.1182	0.0724
	sym1	**sym2**	**sym3**	**sym4**	**sym5**
Linear chirp	0.4414	0.3695	0.3090	0.2484	0.2594
Quadratic chirp	0.4531	0.3217	0.2743	0.2434	0.2266
Logarithmic chirp	0.4238	0.3419	0.2448	0.2451	0.2000
	coif1	**coif2**	**coif3**	**coif4**	**coif5**
Linear chirp	0.3438	0.2798	0.2606	0.2394	0.2256
Quadratic chirp	0.3142	0.2619	0.2593	0.2406	0.2129
Logarithmic chirp	0.3073	0.2440	0.2327	0.2193	0.2044

Wavelet packets with the B-spline bases show the best results in comparison with other orthogonal wavelet packet bases; thus, during WPT, the B-spline wavelet packets detect non-stationary components in the signal most clearly. As observed in Table 1, the values of IDSD for wavelet bases tend to the minimum when their order grows, and the best IDSD values were obtained for B-spline bases. The linear chirp gives higher disturbances in the scalogram, while the logarithmic chirp gives lower ones. Due to the good results obtained for B-spline wavelet packet bases in tests on synthetic signals, they were applied to the real vibration signals in the fault identification problem of rolling bearings.

4. Faults identification in rolling bearings using wavelet packet decomposition

4.1. Problem description and methodology

In the present study, the rolling bearing condition was analysed using several wavelet packets. The tests were performed based on the study presented in [14]. In this study, four ZA-2115 double row bearings manufactured by Rexnord installed on a shaft driven by an AC motor and tested. The bearings

have 16 roller elements in each row, a pitch diameter D_p of 71.5 mm, roller diameter D_r of 8.4 mm, and the contact angle ψ of 15.17°. The rotation frequency f_r of the shaft was constant and equaled 33.33 Hz (2000 rpm). The measurements were carried out on the bearing housing using PCB353B33 accelerometers. The vibration data was collected by a NI DAQ Card 6062E with a sampling rate of 20 kHz. The signals for the present analysis were kept from NASA Ames Prognostics Data Repository [21].

In the present study, the signal from the first bearing with a faulty outer race was analysed. The bearing worked in heavy-load conditions over 160 hours. The characteristic frequency for the outer race defect f_d could be calculated from the following formula [15]:

$$f_d = \frac{nf_r}{2}\left(1 - \frac{D_r}{D_p}\cos\psi\right) \tag{14}$$

Where n is the number of roller elements. In the present case, we obtain the characteristic defect frequency of 293.54 Hz.

The defect detection was performed based on wavelet packet decomposition (WPD) using the following algorithm. The original signals were decomposed on the third level. The level of decomposition in most bearing damage detection problems is assumed as third or fourth [22]. Then, on the assumed decomposition level, the energy vector is determined for all of the sets of wavelet coefficients. The set with the maximum value of energy is considered in further analysis. The energy value is greater when the number of harmonics consisted in a given decomposed realisation is high. For the selected set of wavelet coefficients, the frequency spectrum and the magnitude component with a characteristic defect frequency is extracted and analysed.

4.2. Fault identification and comparative analysis of wavelet packet bases

Based on the presented algorithm, WPD was executed for the signal from faulty bearing using a *bsp5* wavelet packet base. The first set of wavelet coefficients $c_{3,0}$ was omitted because it contains information about low frequencies, which usually represent harmonics from unbalancing, misalignment, etc. [22]. Then, the energy values were determined for the rest of the wavelet coefficient sets. It could be noticed that the set with a higher energy value is $c_{3,2}$ and for this set we determine the spectrum. The energy variation for wavelet coefficient sets on the third level of decomposition and the spectrum for $c_{3,2}$ is presented in Fig. 5. In the spectrum, the component which comes from the fault is clearly identified (the highest peak). Other peaks represent the rotation frequency harmonics.

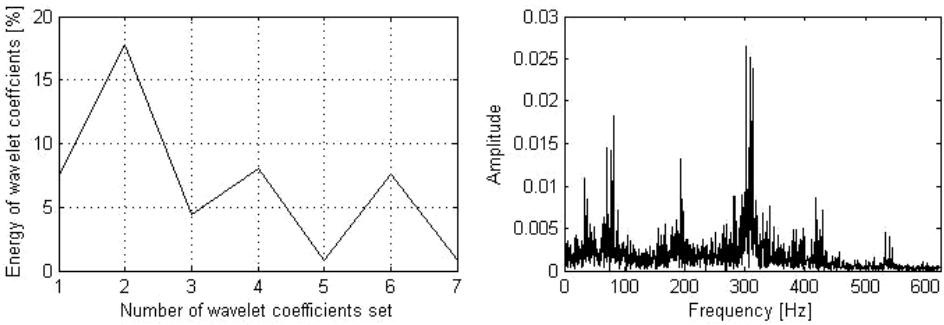


Fig. 5. Energy of wavelet coefficients set on third decomposition level and the spectrum of the best wavelet coefficients set
Rys. 5. Energia zbioru współczynników falkowych na trzecim poziomie dekompozycji oraz widmo dla najlepszego zbioru współczynników falkowych

For the comparison of different wavelet packet bases, we determine the mean value of amplitudes in the spectrum and calculate the ratio R of the fault component amplitude to the mean value. The higher value of R denotes the better sensitivity of a given base. In this analysis, we compare wavelet packet bases used in [12,13,22], some other orthogonal bases and B-spline bases. The results of a comparison are presented in Table 2, where symbols denote the bases of *bior* – biorthogonal wavelet, *rbio* – reverse biorthogonal wavelet, *dmey* – discrete Meyer wavelet.

Table 2. Comparison of different wavelet packet bases sensitivity
Tabela 2. Porównanie czułości różnych baz pakietów falkowych

haar	**db3**	**db5**	**db7**	**db10**	**db12**	**db15**
14.3372	14.5636	12.3543	12.8808	14.4217	14.5438	16.2800
sym2	**sym3**	**sym5**	**sym7**	**sym10**	**sym12**	**sym15**
15.1338	14.5636	12.9303	12.5914	15.4521	14.6641	15.4401
coif1	**coif3**	**coif5**	**bior1.1**	**bior2.4**	**bior3.9**	**bior6.8**
14.8011	15.2622	16.1934	14.3372	12.6599	14.0450	15.1294
rbio1.1	**rbio2.4**	**rbio3.9**	**rbio6.8**	**dmey**	**bsp2**	**bsp3**
14.3372	12.3030	14.2793	16.2204	15.1700	12.7392	16.9109
bsp4	**bsp5**	**bsp7**	**bsp8**	**bsp10**	**bsp12**	**bsp15**
16.8412	16.8598	15.1383	18.4434	18.2720	18.1280	17.9483

The comparison shows that the most sensitive to the fault characteristic frequency are the B-spline bases. The WPD filters constructed based on these wavelets allow the extraction of the harmonics with the best purity.

5. Conclusions

In this paper, the B-spline wavelet packets were constructed and their main properties were presented. Their effectiveness in the signal processing in diagnostic problems was investigated based on the synthetically generated non-stationary signals. The B-spline wavelet packet bases show the best purity of obtained scalograms. Then, they were applied to the fault identification problem of the rolling element bearings using wavelet packet decomposition. The characteristic fault frequency of the bearing was clearly identified. In comparison with other bases, the B-spline bases show the best purity of the identified fault component in the spectrum. The presented research confirms that the B-spline wavelet packets could be used in practical diagnostic problems as well.

References

[1] Timofiejczuk A., Methods of analysis of non-stationary signals, Silesian University of Technology Publishing House, Gliwice, 2004 [in Polish].
[2] Katunin A., Moczulski W., Faults detection in layered composite structures using wavelet transform, Diagnostyka 1(53), 2010, pp. 27-32.
[3] Wysogląd B., Bearing condition diagnostics using entropy of signal in frequency domain, Diagnostyka 2(50), 2009, pp. 65-68.
[4] Douka E., Loutridis S., Trochidis A., Crack identification in beams using wavelet analysis, International Journal of Solids and Structures, 40, pp. 3557-3569, 2003.
[5] Douka E., Loutridis S., Trochidis A., Crack identification in plates using wavelet analysis, Journal of Sound and Vibration, 270, pp. 279-295, 2004.
[6] Katunin A., Korczak A., The possibility of application of B-spline family wavelets in diagnostic signal processing, Acta Mechanica et Automatica, 3(4), pp. 43-48, 2009.
[7] Katunin A., Construction of high-order B-spline wavelets and their decomposition relations for faults detection and localization in composite beams, Acta Mechanica et Automatica, 2010, submitted.
[8] Katunin A., Identification of multiple cracks in composite beams using discrete wavelet transform, Scientific Problems of Machines Operation and Maintenance 45, 2(162), 2010, pp. 41-52.
[9] Coifman R., Meyer Y., Wickerhauser M.V., Size properties of wavelet packets, in: Ruskai M.B. et al. (Eds.), Wavelets and their applications, Jones and Bartlett, Boston, 1992, pp. 453-470.
[10] Misiti M., Misiti Y., Oppenheim G., Poggi J., Wavelet Toolbox®. For use with MATLAB®, The MathWorks Inc., 1996.
[11] Glabisz W., Wavelet packet analysis in mechanics problems, Wroclaw University of Technology Publishing House, Wroclaw, 2004 [in Polish].
[12] Wysogląd B., Rolling element bearing diagnosing with the use of wavelet packet transform, Diagnostyka 30, 2004, pp. 195-198 [in Polish].
[13] Wysogląd B., Application of entropy-based analysis of signals to identification of rolling element bearings faults, Diagnostyka 3(51), 2009, pp. 65-68.
[14] Qiu H., Lee J., Lin J., Yu G., Wavelet filter-based weak signature detection method and its application on rolling element bearings prognostics, Journal of Sound and Vibration 289, 2006, pp. 1066-1090.

[15] Eren L., Devaney M.J., Bearing damage detection via wavelet packet decomposition of stator current, IEEE Transactions on Instrumentation and Measurements 53, 2, 2004, pp. 431-436.
[16] Fan X., Zuo M.J., Gearbox fault detection using Hilbert and wavelet packet transform, Mechanical Systems and Signal Processing 20, 2006, pp. 966-982.
[17] Wei Z., Yam L.H., Cheng L., Detection of internal delamination in multi-layer composites using wavelet packets combined with modal parameter analysis, Composite Structures 64, 2004, pp. 377-387.
[18] Choi S., Detection of valvular heart disorders using wavelet packet decomposition and support vector machine, Expert Systems with Applications 35, 2008, pp. 1679-1687.
[19] Chui C.K., On cardinal spline wavelets, *Wavelets and their applications*, Jones and Bartlett, Boston, 1992, pp. 419-438.
[20] Ueda M., Lodha S., Wavelets: an elementary introduction and examples, University of California, Santa Cruz, 1995.
[21] Lee J., Qiu H., Yu G., Lin J., Rexnord Technical Services, Bearing Data Set, IMS, University of Cincinnati. NASA Ames Prognostics Data Repository, [http://ti.arc.nasa.gov/project/prognostic-data-repository], NASA Ames, Moffett Field, CA, 2007.
[22] Nikolaou N.G., Antoniadis I.A., Application of Wavelet Packets in bearing fault diagnosis, Proceedings of the 5th WSES International Conference on Circuits, Systems, Communications and Computers (CSCC 2001), Rethymno, Greece, 2001, pp. 12-19.

Manuscript received by Editorial Board, February, 2nd, 2011

B-splajnowe pakiety falkowe i ich zastosowanie w wielorozdzielczej analizie sygnałów niestacjonarnych

S t e s z c z e n i e

Wzrastające wymagania do stanu technicznego maszyn powodują rozwój nowych metod diagnostycznych dla detekcji i identyfikacji uszkodzeń w możliwie wczesnej fazie. Większość z tych metod bazuje na analizie sygnałów drganiowych. Klasyczne metody często nie dają pełnej informacji o aktualnym stanie maszyny. Dlatego niezbędny jest rozwój odpowiednich metod diagnostycznych. Niektórymi z obiecujących metod analizy sygnałów są metody oparte na transformacji falkowej dające możliwość efektywej diagnostyki niestacjonarnych sygnałów drganiowych w dziedzinie czasowo-skalowej. Uogólnieniem transformacji falkowej jest pakietowa transformacja falkowa pozwalająca wydobyć z sygnału dodatkowe korzystne cechy. Jednak efektywność w przypadku metod diagnostycznych bazujących na przekształceniu falkowym jest zdeterminowana wyborem odpowiedniej funkcji falkowej. W niniejszej pracy autor wprowadza nowe pakiety falkowe bazujące na falkach B-splajnowych. Analiza porównawcza ich efektywności była przeprowadzona na niestacjonarnych sztucznych sygnałach. B-splajnowe pakiety falkowe były zastosowane do oceny stanu łożysk tocznych.