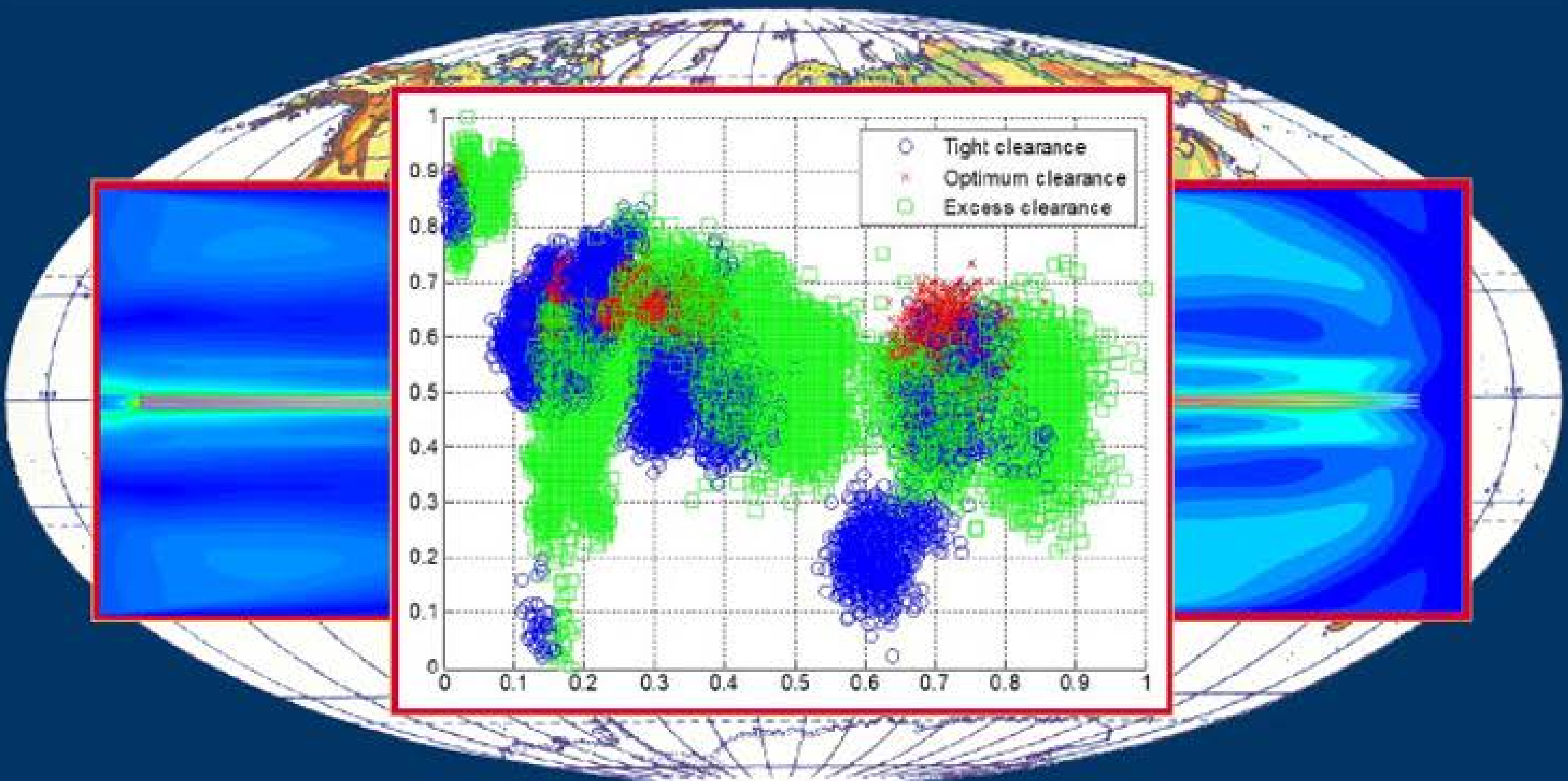
Vol. 23. No 2, 2021

ISSN 1507-2711
Cena: 25 zł

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY



Polskie Naukowo Techniczne Towarzystwo Eksploatacyjne
Warszawa

Polish Maintenance Society
Warsaw

Indexed by:
Web of Science Group

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

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

Ministry of Science and Higher Education
Republic of Poland

Task *„Implementation of procedures ensuring the originality of scientific papers published in the quarterly „Eksploatacja i Niezawodność – Maintenance and Reliability"* financed under contract 532/P-DUN/2018 from the funds of the Minister of Science and Higher Education for science dissemination activities.

All the scientific articles have received two positive reviews from independent reviewers.
Our 2018 Impact Factor is 1.525

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

INFORMATION FOR AUTHORS

Terms and Conditions of Publication

1. The quarterly „Eksploatacja i Niezawodnosc - Maintenance and Reliability" publishes original papers written in English. Authors are strongly recommended to have their manuscripts checked by a native English speaker or a professional language editing service before submission, in order to ensure that the language is acceptable. Submissions may be rejected due to low quality of the translation.
2. The Quarterly welcomes all original submissions of articles that comply with its Aims and Scope. If a paper contains any material that has already been published (e.g. figures), then this must be declared by referencing the source in the manuscript, and authors are responsible for obtaining relevant permissions to republish. Evidence of permission granted must be available upon request.
3. All submissions should be made online via the peer-review system. Authors can use the „Articles" button to make a new submission or check the status of their manuscript. New users will need to create an account.
4. Submissions are made on the understanding that the manuscript has not been published elsewhere and is not currently under consideration by another journal published by any other publisher.
5. The submitting author is responsible for ensuring that the submission has been approved for publication by all co-authors. Submitting authors' responsibilities include especially the following issues: plagiarism, authorship, conflict of interest and misleading or misreporting of findings. In all cases, the Editorial Office reserves the rights to contact the authors' institutions if serious misconduct has been identified, and to retract any articles in which serious misconduct has been identified.
6. Each submission is checked for suitability when received by the Editorial Office, and may be rejected without review if it is outside the scope of the journal, is obviously of insufficient quality, or is missing important sections. This ends the assessment procedure. Re-submission of a rejected manuscript is not possible.
7. Reviewers are selected in accordance with the principle of avoiding conflict of interest (examples of conflicts of interest include a personal relationship between reviewer and author, occupational subordination, direct scientific cooperation over the last two years prior to the review). The journal operates a single-blind peer review system.
8. The review is provided in written form and concludes with an explicit recommendation to accept, revise or reject the paper for publication. If a negative review is given, the submission is rejected, which ends the assessment procedure. Re-submission of a rejected manuscript is not possible.
9. The Editor-in-Chief makes the publication decision after receiving at least two external reviewer reports with recommendations. On receiving the required number of reviews, the Editors pass them immediately to authors. The Editors require that authors prepare answers to reviewers' comments, which are then sent to the referees. The Editor-in-Chief will make a decision to accept, accept with minor revisions, accept with major revisions, or reject the paper for publication. The reasons for the decision will be communicated to the authors. When the decision of minor/major revisions is made, and the authors do not revise their articles satisfactorily after receiving reviewer reports, then the Editor-in-Chief reserves the right to reject the article. When revised articles are received, they will either be sent out for further review or the Editor-in-Chief will make a decision depending on the level of revision requested.
10. The time required for reviewing a manuscript and making a publication decision may vary a lot from one submission to the next since it is sometimes difficult to find suitable reviewers, and there may be delays in receiving reviewer reports. The Editor-in-Chief and the Editorial Office try their best to minimize the time from submission to first decision. The journal aims to make the first decision within 7 days after the Editorial Office check and the second decision within 30 days after the review, but no guarantees can be made in this regard.
11. Submissions with positive reviews are subjected to an originality check. Only those articles that have successfully passed the anti-plagiarism test are finally accepted for publication.
12. All accepted articles are published as open-access and distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Full texts of all printed articles are available on the website of the Quarterly.

Fees

The publication fee for one text is 410 EUR (1700 PLN for submissions from Polish institutions).

Technical requirements

1. All papers must be submitted in a Microsoft Word document format. Drawings and photos should be submitted as separate graphic files in *.tif, *.jpg or *.cdr format.
2. A manuscript should include:

   Names of authors, **title**, **abstract**, and **keywords** that should complement the title and abstract. **The corresponding author** should be indicated. Abstract should not exceed 1100 characters (including spaces) and **must be divided into the following sections**: Objectives, Methods, Results, and Conclusions.

   Highlights (a short collection of bullet points that convey the core findings and provide researchers with a quick overview of the article in text form) - three to five bullet points should describe the essence of the study (e.g. results or conclusions) and highlight what is distinctive about it. Highlights should not exceed 90 characters (including spaces) in each bullet point.

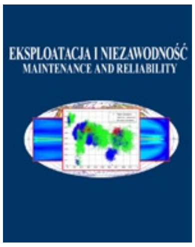
   The text in English with a clear division into sections (please, do not divide words in the text).

   Tables, drawings, graphs, and photos included in the text should have descriptive English-language captions.

   If this can be avoided, **no formulae and symbols should be inserted into text paragraphs by means of a formula editor**.

   References (written in accordance with the enclosed reference format). Authors using the Mendeley and the Zotero citation plugins and styles support may download the required format file from the repository.

   Author data should include first names and surnames along with scientific titles, affiliation, detailed postal address and official e-mail address.



Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Jasiulewicz–Kaczmarek M, Antosz K. Żywica P, Mazurkiewicz D, Sun B, Ren Y. Framework of machine criticality assessment with criteria interactions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 207–220, http://doi.org/10.17531/ein.2021.2.1.

Framework of machine criticality assessment with criteria interactions

Indexed by: Web of Science Group

Małgorzata Jasiulewicz–Kaczmarek [a], Katarzyna Antosz [b], Patryk Żywica [c], Dariusz Mazurkiewicz [d], Bo Sun [e], Yi Ren [e]

[a] Poznan University of Technology, Faculty of Management Engineering, ul. Prof. Rychlewskiego 2, 60-965 Poznan, Poland
[b] Rzeszow University of Technology, Faculty of Mechanical Engineering and Aeronautics, ul. Powstańców Warszawy 8, 35-959 Rzeszów, Poland
[c] Adam Mickiewicz University, Faculty of Mathematics and Computer Science, ul. Uniwersytetu Poznańskiego 4, 61-614, Poznan, Poland
[d] Lublin University of Technology, Mechanical Engineering Faculty, ul. Nadbystrzycka 36, 20-618 Lublin, Poland
[e] School of Reliability and Systems Engineering, Beihang University (BUAA), No.37 XueYuan RD. Haidian, Beijing 100191, China

Highlights

- A review of machines criticality assessment criteria was presented.
- A novel model of a machine criticality assessment is proposed.
- The model combines the importance of the machine criticality assessment criteria with interactions between them.
- The machine criticality assessment model for the aviation industry is presented.

Abstract

Criticality is considered as a fundamental category of production planning, maintenance process planning and management. The criticality assessment of machines and devices can be a structured set of activities allowing to identify failures which have the greatest potential impact on the company's business goals. It can be also used to define maintenance strategies, investment strategies and development plans, assisting the company in prioritizing their allocations of financial resources to those machines and devices that are critical in accordance with the predefined business criteria. In a criticality assessment process many different and interacting criteria have to be taken into consideration, despite the fact that there is a high level of uncertainty related to various parameters. In addition, not all assessment criteria are equally important. Therefore, it is necessary to determine the weight of each criterion taking into account different requirements of machine criticality process stakeholders. That is why a novel model of a machine criticality assessment is proposed in this paper. The model extends the existing methods of assessing machines criticality, taking into account not only the importance of machine criticality assessment criteria, but also possible interactions between them.

Keywords

machine criticality assessment, assessment criteria, assessment methods, interactions.

1. Introduction

The recent rapid development of production systems related to automatization and digitalization has required a new approach to defining the function and role of a technical object in a production processes. Due to the client's requirements for a product, the technologies used for their realization and the impact of failures on people and natural environment, the companies not only must plan maintenance activities, but also have to define the priorities for implementations of these activities taking into account the role they play in business goals [13, 41, 75]. Therefore, an important issue for any company is a machine criticality assessment. Referring to [1, 4, 70] criticality is a fundamental category of the production and maintenance process planning and management. A machine and device criticality assessment is a structured set of activities that allows to identify machines and devices which failures have the greatest potential impact on the company's business goals. It can be used to define maintenance strategies, investment strategies and development plans, assisting the company in prioritizing the allocation of financial resources to those machines and devices that are critical in accordance with the predefined business criteria [27, 37, 47, 60, 89]. Moreover, according to Roy [70], prioritizing of maintenance activities eliminates their instability and variability in activities, thereby increasing resource efficiency and reducing maintenance costs.

Although the literature review describes many methods for assessing the machine criticality and decision making systems in this area, itis still not a simple task [6, 13, 40, 48]. First of all, in a machine criticality assessment process many different and interacting criteria have to be taken into consideration [21]. Secondly, due to the quality and method of data acquisition there is a high level of uncertainty related to various parameters such as: time between failures, time to repair and the quantity of spare parts needed for a repair [46]. Thirdly, not all assessment criteria are equally important. Therefore, it is neces-

E-mail addresses: M. Jasiulewicz–Kaczmarek - malgorzata.jasiulewicz-kaczmarek@put.poznan.pl, K. Antosz - katarzyna.antosz@prz.edu.pl, P. Żywica - bikol@amu.edu.pl, D. Mazurkiewicz - d.mazurkiewicz@pollub.pl, B. Sun - sunbo@buaa.edu.cn, Y. Ren - renyi@buaa.edu.cn

sary to determine the weight of each criterion taking into account different requirements of machine criticality process stakeholders [75]. Considering the above issues, a novel model of a machine criticality assessment is proposed in this paper. The proposed model extends the methods of assessing the machine criticality described in the literature and used in practice taking into account not only the importance of machine criticality assessment criteria, but also the interactions between them.

This paper is organized as follows: in Section 2 the literature review according the criteria and methods used for a machine criticality assessment is presented. Then, in Section 3 a novel framework of a machine criticality assessment is developed. Moreover, in this section the study results of the importance for the criticality assessment in different industries are presented. In Section 4 a machine criticality assessment model for the aviation industry is presented. Finally, the conclusions and direction of the future research are presented.

2. Problem statement of a machine criticality assessment

2.1. Criticality criteria

As mentioned in work [80], criticality is a measure of importance defined on the basis of the analyzed factors. Moreover, criticality is used as a comparative measure to assess the consequences of actions taken and it can be used as a measure to highlight the differences between individual machines and action scenarios (action strategies). The criteria adopted for the evaluation may affect the final criticality often differ from one organization to another. They are often dependent on the type of assets (resources) as well as adopted rules in the organization. In the literature many criteria for assessing the criticality of machines are defined. Because the classification of the criteria for assessing the criticality of machines proposed in the literature is not unambiguous and may cause problems related to their interpretation two – level- hierarchical classification for the machine criticality assessment criteria was proposed in Table 1.

Moreover, in the literature in different areas, other criteria of machine criticality assessment are proposed. In the area of manufacturing systems, the following factors are indicated: redundancy, work load [4], production integrity [18, 50, 63, 86], machine importance for a process [5], breakdown time and stability of the machine [75], sensitivity of operation [36], bottleneck and impact on throughput [44], applicability of CBM [63] and reliability [54, 63].

Furthermore, in the medical assessment, the following criteria are used: risk, user competence and standards [71], performance assurance [14, 71], support availability, clinical acceptability [14], function [14, 80], recalls and hazard alerts and maintenance requirements [80].

In the oil refinery assets the following criteria are proposed: failure detection and failure severity [33]. What's more, the customer's inconvenience criterion [74] and effect of power generation in thermal power plants equipment assessment [34] as well as the impact of business (shutdown duration) in chemical plants equipment assessment [65] are proposed.

2.2. Criticality method assessment used and criticality levels

There are many different methods presented in the literature for assessing the criticality of machines. These methods use a variety of evaluation criteria and are used in different industries. The most

Table 1. Criteria for assessing the criticality of machines – literature survey

No	Main criteria	Sub - criteria	References
1.	Safety	Degree of influence on working conditions. Machine failure costs due to health, safety and environment.	[5, 18, 24, 26, 34, 38, 50, 53, 54, 63, 64, 65, 66, 67, 75, 86]
2.	Environment	In case of failure, the degree of risk for the environment. Working environment.	[26, 28, 34, 53, 54, 64, 65, 74, 86]
3.	Maintainability	MTTR (Mean Time To Repair). OEE (Overall Equipment Effectiveness). Failure detection. Failure frequency. Failure severity. Downtime length.	[2, 5, 16, 19, 20, 31, 33, 41, 43, 53, 63, 6, 73, 75, 79, 92]
4.	Quality	The degree of influence of a machine on the implementation of other operations in case of a failure. Number of nonconformities due to the machine failure during the year. The degree of influence of a technical object on the quality of the final product. Costs of non-conforming products as a result of a failure.	[4, 18, 19, 20, 26, 31, 50, 54, 66, 67, 68, 90]
5.	Age	Age of machine.	[4, 14, 68, 80]
6.	Cost	Costs incurred by production for machine downtime (breakdown). Costs of non-conforming products as a result of failure. Failure elimination costs (excluding OHS and environmental costs). Machine failure costs due to health, safety and environment.	[4, 5, 6, 14, 19, 20, 26, 31, 33, 53, 65, 68, 72, 74]
7.	Risk	Mission criticality. Operating conditions and equipment accidents.	[19, 20, 28, 31, 80]
8	Availability	Average downtime of a technical facility due to failures and repairs. Availability of the required personnel. MTBF (Mean Time Between Failures). Machine replacement in case of failure (Machine changeability). Work load.	[2, 31, 43, 53, 63, 66, 67, 76]
9.	Spare parts	Spare parts availability.	[36, 63]

commonly used method is the AHP (Analytical Hierarchy Process) method. This method is used to select the best alternative and analyze possible alternatives [17, 76, 84]. In the work [5] the authors indicate that thanks to the great alternatives it can be used to arrange a large number of machines. In order to create the ranking, it is possible to integrate both qualitative and quantitative criteria, as well as their integration [14, 46, 69]. The AHP method has been successfully applied to the classification of equipment in the thermal power plant [34, 74], for the prioritization in the medical industry [80] and machinery classification in the plastics processing industry [1]. Moreover, the authors in [76] noted that it can be integrated with other methodologies, such as the Delphi method. This method relies heavily on experts. Such integration into the classification of equipment in an oil refinery is presented in [33]. In [85] it was noted that the calculation of the coefficients of coherence makes this method more reliable compared to the additive weighting method presented in [31, 75].

Additionally, in order to evaluate the criticality of machines, the rules of interconnection with the fuzzy logic and fuzzy grouping are used. The result of this assessment is the identification of different categories of machines. Rules are an appropriate method to evaluate a large number of machines for which common strategies and operating procedures are defined [41, 54]. The combination of rules with the fuzzy logic is presented in [36]. Furthermore, [28] presents a fuzzy cluster analysis structure divided into four sub-hierarchy models. Fuzzy grouping is also presented in the literature as a possible method for assessing the importance or criticality of equipment. Guo et al. [26] found that fuzzy assessments can deal with imprecise information better, which can be beneficial for companies. However, its effective application depends on the function of membership and a set of weighting factors. The fuzzy logic application has some advantages. However, it is a complex methodology and difficult to advance as it requires some simulations before use [85].

Moreover, in the literature, the FMEA method (Failure Mode And Effects Analysis) is used to assess the criticality of machines. It allows the identification of the factors that were taken into account for the criticality assessment. Most often this method is used to assess the types of failures, with particular emphasis on the likelihood of failures and their consequences, taking into account such factors as: redundancy, use, quality, age of machine and costs [4, 68]. Another variant of the FMEA – FMCEA (Failure Mode, Effects, And Criticality Analysis) method takes into account additional factors such as environmental aspects when assessing criticality [12]. Based on the value of RPN (Risk Priority Number) index and risk matrix the machine criticality and the maintenance strategy are determined.

The applied criticality method allows to determine the machine criticality level (machine category). The authors in the works [8, 29, 57, 82] classified machines into three groups on the basis of the ABC (Activity Based Classification) analysis method. The main goal of this method is based on Pareto's principle, which classifies the top 15–20% goods occupied 65–80% value of the whole system into A group, the following 30–40% goods occupied 15–20% value of the whole system into B group, and the other 40–55% goods occupied 5–15% value of the whole system into C group [29]. Additionally, a scoring system is used to assess the criticality [68, 75]. The authors of the mentioned works used ABC or ABCD classification levels. ABC – type classifies machines into three groups: category A – machines which need special control, category B – machines which need less control and category C – machines which do not need any special control. The ABCD - type classification defines four categories of machines as: category A – particularly important (bottleneck), category B – important, category C – relatively important, category D – not used.

In addition to the ABC classification, there is another method of assessing the criticality of machines, that is the GUT (Gravity, Urgency and Tendency) matrix [15].

2.3. Research challenges

The analysis of the literature showed that many criteria are proposed for assessing the criticality of machines. The same criteria are often defined differently, e.g. breakdown /failure. Moreover, sometimes it is difficult to understand the meaning of the criterion unequivocally on the basis of the description provided by the author. The criteria analysis presented in Table 1 allowed to systematize them. Nevertheless, there are no unequivocal studies indicating the real usefulness and importance of the specific criteria for different industries.

In addition, various methods are available for assessing the criticality. These methods take into account various criteria proposed (as discussed above). However, these criteria are often analyzed independently, or their dependence is analyzed to a little extent. In the literature, the method that takes into account the interactions between the individual evaluation criteria used to assess the criticality of machines is not described. From the point of view of the machine criticality assessment, most of the proposed criteria aggregation methods have some drawbacks. Namely, they do not reflect the interaction between the criteria. In the real manufacturing environment machine criticality criteria are usually not independent (there are some interactions among the criteria, positive or negative effects between them) Thus, an appropriate function must be used to aggregate multiple information sources and to handle an interactive relationship. An example of such an environment is a noisy environment where complex criterion relationships between worker and machine can be identified [61]. Since ignoring the interaction between the assessment criteria may lead to distortions of its outcome and, consequently, ineffective and inefficient decisions, commonly used aggregation and ranking methods such as AHP, SAW and WSAW do not apply in this problem.

That is why, in this paper, a new method of machine criticality assessment is proposed.

As part of the work, the following research questions were taken into consideration:

1. Is there a difference in the perception of the criteria importance of the machine criticality assessment in various industries?
2. Which method is able to model the importance of machine criticality assessment criteria and the interaction between them?
3. How important are the particular machine criticality assessment criteria and the interactions between them in the case study industry?

3. The framework of machine criticality assessment

3.1. Development of the methodology

Machine criticality is a complex concept and depends on many factors. "Intuition" is usually not sufficient to make an objective decision about which machine is important and which is not. It is necessary to build a structured method to support decision makers in the machine criticality assessment process. A general scheme of the machine criticality assessment method used in this paper includes three main stages: (1) Selection of criteria, (2) Criteria assessment, and (3) Selection of the appropriate aggregation function (Figure 1).

Based on the final Machine Criticality Index (MCI), it is possible to define the prioritization of maintenance actions in order to ensure that the production system works as close to its nominal capacity as possible.

3.2. Identification of criticality assessment criteria

The main standard for evaluating the machine criticality is the criteria. Based on the analysis of the literature (chapter 2.1), 24 criteria most frequently matched and used to assess the criticality of machines were selected: C1 - Operators' competences; C2 - Machine replacement in case of failure; C3 - Degree of influence on other tasks in case of a failure; C4 - Costs incurred by production for machine downtime (breakdown); C5 - Number of nonconformities due to a machine failure during the year; C6 – Degree of machine influence on the final

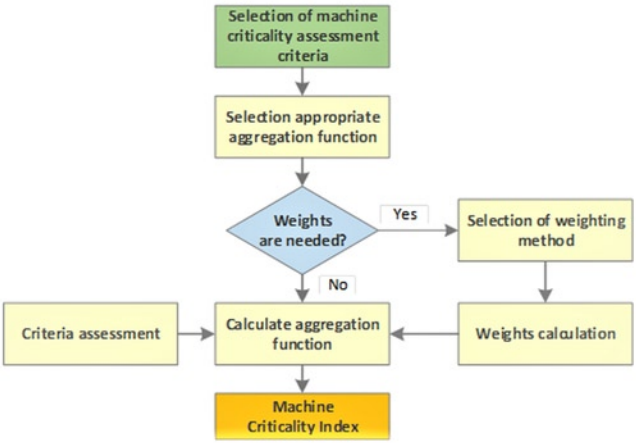


Fig. 1. A generic process for the machine criticality assessment

product quality; C7 - Costs of non-conforming products as the result of a failure; C8 - Frequency of failures per year; C9 - Average downtime of a technical facility due to failures and repairs; C10 – Spare parts availability; C11 - Failure elimination costs (excluding OHS and environmental costs); C12 - In case of a failure, the degree of influence on working environment; C13 - In case of a failure, the degree of risk to environment; C14 - Machine failure costs due to health, safety and environment; C15 - Age of the machine; C16 - OEE; C17 - MTBF; C18 - MTTR; C19 - Failure severity; C20 - Failure detection; C21 - Customer's inconvenience; C22 - Mission criticality; C23 - Operating condition; C24 - Work load.

In the next step, these criteria were assessed by company experts in order to identify the most important criteria from the industry point of view. The research on the perception of the importance of the machine criteria criticality assessment was carried out at the turn of 2019 in 2020 in small, medium-sized and large production companies of various industries selected for the research purposely. A group of 66 production companies participated in the study, of which 22.39% were enterprises from the automotive industry, 29.85% from the food industry, 31.34% from the aviation industry and 16.42% from other industries, e.g. medical, furniture, railway, printing house, etc. The biggest group of them was the large enterprises 57.58%, and smallest group was small sized companies (6.06%) (Figure 2).

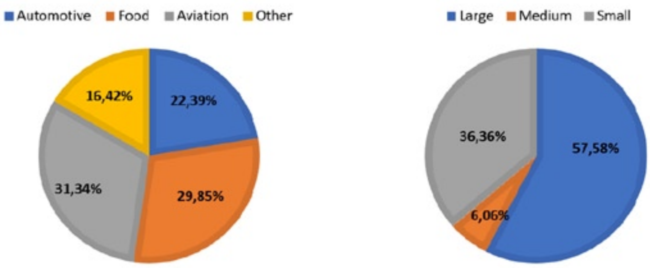


Fig. 2. Structure of the enterprises participating in the survey

The survey was conducted with experts from these enterprises. The experts were asked to determine the degree of importance of the 24 criteria on a scale from 1 to 5, where 1 meant – not important, while 5 very important. The data set obtained from the enterprises was subjected to a statistical assessment (an average assessment value - $\overline{X}$, a standard deviation – S, $\sum$ - total importance obtained by the criterion in a given industry) (Table 2).

Analyzing the results presented in Table 1 there are visible differences not only in the assessment of the importance of the criteria in individual industries, but also in each individual criterion in a given industry. In case of the automotive industry, the highest compliance in the assessment of the criterion by enterprises was identified for the C5 criterion - the value of a standard deviation is s = 0.408. The lowest compliance was noted for the C16 criterion (s = 1.188). In the food and aviation industries, the highest compliance in the assessment of the importance was achieved by the C6 criterion, with the following values of a standard deviation - food s = 0.413, aviation s = 0.359. On the other hand, the lowest compliance was achieved by the C17 criterion in the food industry (s = 1.223), and the C24 criterion in the aviation industry (s = 0.949). Additionally, it should be noted that in case of the aviation industry, the values of the standard deviation for the assessed criteria were the lowest, what proves high consistency in assessing the importance of criteria in individual companies in this industry. In case of enterprises identified as "other", the highest compliance in the assessment of the criterion by the enterprises was identified for the C8 criterion - the standard deviation value is s = 0.601, and the lowest, the same as in the automotive industry, for C16 (s = 1.481). An average value of the importance obtained by the analyzed criteria for individual industries is presented in Figure 3.

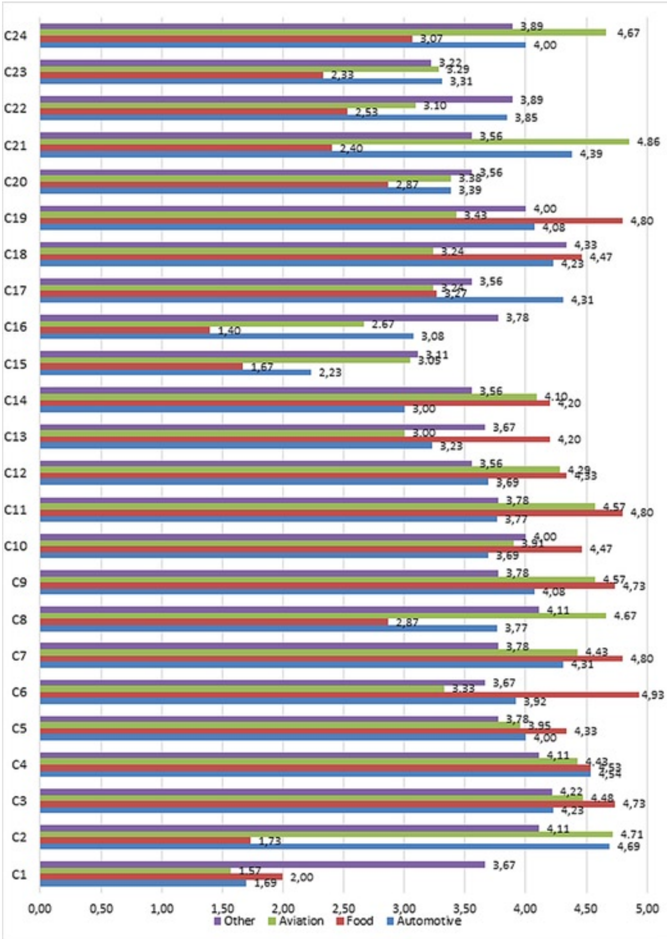


Fig. 3. An average assessment value of importance ($\overline{X}$) obtained by the analyzed criteria for individual industries

The analysis of the above results (Figure 3) made it possible to identify common assessment criteria for individual industries. When identifying the common criterion, similar or insignificant differences (± 0.5) in the obtained average assessment value ($\overline{X}$) for a particular criterion were taken into account. The criteria common for all industries are: C3, C4 and C5. However, the C1 and C7 criterion is common for the automotive, aviation and food industries. Criteria C2, C20, C22 and C23 are common for the aviation, automotive industries and the enterprises defined as "other". The C6 and C10 criterion obtained the highest value (4.93 and 4.47) for the food industry, but this criterion is common for aviation, automotive industries and the enterprises defined as "other". On the other hand, the C8, C15 and C16 criterion is common for the aviation industry and the enterprises defined as "other". Moreover, the C9 criterion obtained similar $\overline{X}$ values for

Table 2. Perception of machine criticality factors importance

Criteria	Automotive			Food			Aviation			Other		
	Σ	$\bar{X}$	s	Σ	$\bar{X}$	s	Σ	$\bar{X}$	s	Σ	$\bar{X}$	s
C1	22	1.692	0.480	30	2.000	0.655	33	1.571	0.676	33	3.667	1.414
C2	61	4.692	0.480	26	1.733	0.884	99	4.714	0.561	37	4.111	1.364
C3	55	4.231	0.725	71	4.733	0.458	94	4.476	0.814	38	4.222	0.972
C4	59	4.538	0.519	68	4.533	0.743	93	4.429	0.507	37	4.111	0.782
C5	52	4.000	0.408	65	4.333	0.900	83	3.952	0.669	34	3.778	0.833
C6	51	3.923	0.494	74	4.933	0.413	102	3.333	0.359	33	3.667	1.225
C7	56	4.308	0.630	72	4.800	0.414	93	4.429	0.598	34	3.778	1.394
C8	49	3.769	0.439	43	2.867	1.060	98	4.667	0.483	37	4.111	0.601
C9	53	4.077	0.760	71	4.733	0.458	96	4.571	0.507	34	3.778	0.667
C10	48	3.692	0.751	67	4.467	0.834	82	3.905	0.436	36	4.000	1.323
C11	49	3.769	0.439	72	4.800	0.414	96	4.571	0.507	34	3.778	0.833
C12	48	3.692	1.032	65	4.333	0.976	90	4.286	0.644	32	3.556	1.130
C13	42	3.231	1.013	63	4.200	0.941	98	3.000	0.483	33	3.667	1.118
C14	39	3.000	0.816	63	4.200	0.941	86	4.095	0.539	32	3.556	1.333
C15	29	2.231	0.599	25	1.667	0.488	64	3.048	0.590	28	3.111	0.782
C16	40	3.077	1.188	21	1.400	0.507	56	2.667	0.913	34	3.778	1.481
C17	56	4.308	1.109	49	3.267	1.223	68	3.238	0.539	32	3.556	1.130
C18	55	4.231	1.092	67	4.467	0.915	68	3.238	0.436	39	4.333	1.000
C19	53	4.077	0.641	72	4.800	0.414	72	3.429	0.598	36	4.000	1.000
C20	44	3.385	0.768	43	2.867	0.990	71	3.381	0.865	32	3.556	1.236
C21	57	4.385	0.650	36	2.400	0.507	70	4.857	0.577	32	3.556	0.882
C22	50	3.846	0.376	38	2.533	0.834	65	3.095	0.700	35	3.889	1.054
C23	43	3.308	0.630	35	2.333	0.617	69	3.286	0.463	29	3.222	0.833
C24	52	4.000	0.577	46	3.067	0.594	63	4.667	0.949	35	3.889	1.054
Legend:	(green)	the lowest value of a standard deviation (s) for every industry										
	(red)	the highest value of a standard deviation (s) for every industry										

the automotive industry and the enterprises defined as “other”, with the values of 3.78 and 4.06, respectively. Moreover, it should be noted that this criterion obtained also similar values for the food and aviation industries, 4.73 and 4.57, respectively. Additionally, the criteria C11, C12 and C14 obtained the highest values for the food and aviation industries. The C21 and C24 criteria for the aviation industry achieved the highest values, respectively 4.86 and 4.67. On the other hand, the C13 criterion with the obtained average value of 4.20 and C19 with the average value 4.80 dominated in the food industry. This criterion obtained similar values for the aviation (3.00) and automotive (3.23) industries. The C17 criterion dominated in the automotive industry (with the average value of 4.31), but this criterion is common for the aviation, food industries and the enterprises defined as “other”.

In Figure 4, the total importance obtained by the analyzed criteria for individual industries is presented. The presented results show which of the individual criteria obtained the highest total value for individual industries. In the aviation and automotive industries, the lowest value was obtained by the C1 criterion (value 33, 22). However, in the aviation industry, the highest value was obtained by the C6 criterion, which was also dominant in the food industry. The most important criterion in the automotive industry was the C2 criterion. A completely different situation can be noticed in the enterprises defined as “other”. In this case, the C23 criterion had the lowest value and the C18 criterion the highest. What is more, it should be noted that the criteria from C2 to C14 in the aviation industry obtained the highest values among the surveyed criteria. On one hand, it is justified by the fact that most companies in this industry participated in the research. On the other hand, it turned out that these criteria are the most important from the point of view of machine criticality in this industry.

The criteria from C2 to C14 in the aviation industry were selected for a further analysis (model building). Additionally, the factors determining the choice of these criteria were: the largest number of companies in the aviation industry in the conducted research as well as the fact that in case of this industry, the values of the standard deviation (s) for the assessed criteria were the lowest, what proves a high consistency in assessing the importance of criteria in individual companies (conformity of the assessment).

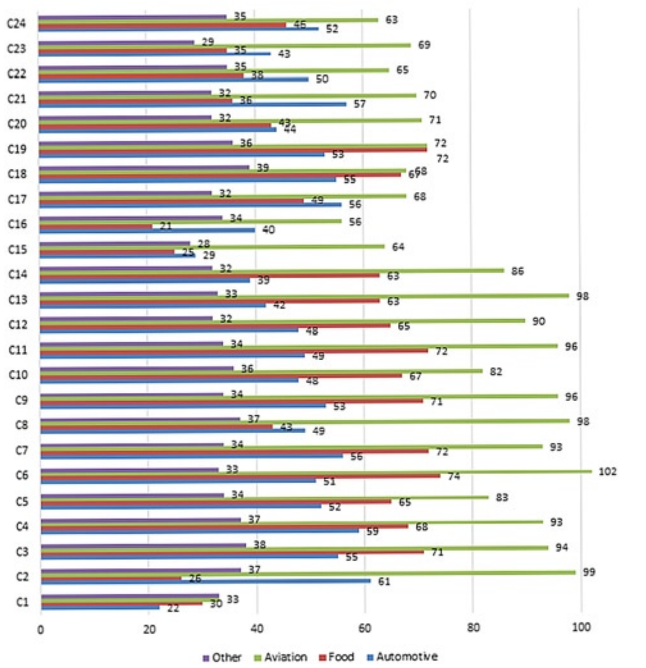


Fig. 4. The total importance obtained by the analyzed criteria for individual industries

After identifying the criteria that will be further analyzed, we can proceed to the next step, i.e. building a function that aggregates individual criteria values into one synthetic result.

3.3. Development of an aggregation function

3.3.1. Assumptions

The purpose of aggregation functions is to combine multiple numerical inputs into a single numerical value, which in some sense represents all the inputs. According to [3] many aggregation functions present some drawbacks, mainly from their natural assumption that input criteria are independent of each other (arithmetic mean, weighted mean, median, mode etc.) and none is capable to find an interaction between criteria [51]. However, this fact does not limit their usability in many complex areas of application [62, 87].

Because the criteria of the machine criticality assessment are usually not independent (there are some interactions among criteria, positive or negative effects among them), an appropriate function must be used to aggregate multiple information sources and to handle an interactive relationship.Based on the literature review [11, 22, 45], in order to solve the problem of aggregating the criteria that are interdependent, a non-additive function that defines a weight, not only for each criterion but also for each subset of criteria, is needed. Thus, these non-additive functions can model both the importance of criteria and the positive and negative synergies between them. Taking the above into account, we propose the use of the machine criticality index (MCI) λ-fuzzy measure and Choquet fuzzy integral, which can handle both the challenges. According to [52] the Choquet integral has good properties for aggregation. It is continuous, non-decreasing, comprised between min and max, stable under the same transformations of interval scales in the sense of the theory of measurement, and it coincides with a weighted arithmetic mean when the fuzzy measure is additive. In view of the characteristics of the Choquet integral, it has been widely applied to multiple attribute decision-making in many areas [7, 9, 22, 23, 35, 56, 78]. However, the interest in the fuzzy integral is mainly due to its ability to represent interactions between criteria . This is due to the fact that weights in a fuzzy measure are assigned to every subset of all criteria.

3.2. Definitions and notations

The fuzzy set theory has been applied to many problems in different fields of science and engineering. In order to describe this theory, some definitions are presented as follows:

Let $X = \{x_1,\ldots,x_n\}$ be the set of all criteria and $\mathcal{P}(X)$ the power set of X.

Definition 1 (Fuzzy measure, [10])**:** A discrete fuzzy measure on X is a set function μ: $\mathcal{P}(X) \rightarrow [0, 1]$ satisfying the following conditions:

1. μ(∅)=0, μ(X)=1 (boundary condition)
2. if $A \subseteq B \subseteq X$ then $\mu(A) \le \mu(B)$ (monotonic condition).

In this context, μ(A) represents the degree of importance of a given criteria set A. This way, additionally to the weight of a single criterion, the weight of an arbitrary criteria combination is also directly described. The fuzzy measure is additive when if $A \cap B = \varnothing$ then $\mu(A \cup B) = \mu(A) + \mu(B)$ and superadditive (subsdditive) when $\mu(A \cup B) > \mu(A) + \mu(B)$ (respectively $\mu(A \cup B) < \mu(A) + \mu(B)$).

Definition 2 (Discrete Sugeno λ-measure)**:** A discrete fuzzy measure is called Sugeno λ-measure if it satisfies:

3. If $A \cap B = \varnothing$, then $\mu_\lambda(A \cup B) = \mu_\lambda(A) + \mu_\lambda(B) + \lambda\mu_\lambda(A)\mu_\lambda(B)$.

Note that (1) and (2) are fundamental properties for any types of a fuzzy measure and (3) is an additional property of λ-measure. To differentiate this measure from other fuzzy measures, λ-fuzzy measure is denoted by μ_λ. Sugeno [77] proved that given those 3 axioms, a fuzzy measure can be uniquely determined using only n = |X| coefficients μ_i that are often called fuzzy densities which represent the degree of importance of the criteria *i-th* and can be calculated with parametric or nonparametric methods. The λ-measure can be calculated using the following formula:

$$\mu_\lambda(\{x_1, x_2,\ldots, x_n\}) = \left|\sum_{i=1}^{n}\mu_i + \lambda\sum_{i_1=1}^{n-1}\sum_{i_2=i_1+1}^{n}\mu_{i_1}\mu_{i_2} + \lambda^{n-1}\mu_1\mu_2\cdots\mu_n\right| = \frac{1}{\lambda}\left|\prod_{i=1}^{n}(1+\lambda\mu_i) - 1\right|$$

Based on the boundary condition in Eq. (1), $\mu_\lambda(X)=1$, λ can be uniquely determined via the following equation:

$$\lambda + 1 = \prod_{i=1}^{n}(\lambda\mu_i + 1), \tag{2}$$

where $\mu_i = \mu(\{x_i\})$, i =1, 2, …, n is known as the fuzzy density function of a single element (singleton), $x_i \in X$.

According to Gürbüz et al. [30] and Hu and Chen [32]:

– If λ < 0 then it implies that the attributes share a redundancy effect. This means a significant increase in the performance of the target can be achieved by only enhancing some attributes in X which have higher individual importance.
– If λ > 0 then it interprets that the attributes share a synergy support effect. This means a significant increase in the performance of the target only can be achieved by simultaneously enhancing all the attributes in X, regardless of their individual importance.
– If λ = 0, then it indicates that the attributes are non-interactive.

Definition 3 (Discrete Choquet integral)**:** Let μ be a discrete fuzzy measure on X. The discrete Choquet integral of function $f: X \rightarrow [0,1]$ with respect to the fuzzy measure μ is defined by:

$$C_\mu(f_1, f_2,\ldots,f_n) = \sum_{i=1}^{n}(f_{(i)} - f_{(i-1)})\mu(A_{(i)}), \tag{3}$$

where: $f_{(i)}$ indicates that the indices have been permutated so that $0 \le f_{(1)} \le \ldots \le f_{(n)} \le 1$, $A_{(i)} = \{x_{(i)}, \ldots, x_{(n)}\}$ and $f_i = f(x_i)$.

Definition 4 (Shapley value - ν_i, [59]: Let μ be a fuzzy measure on X. The Shapley value (or the importance index) for every element $x_i \in X$ is defined by the following formula:

$$\nu_i = \sum_{A \subset X \setminus \{x_i\}}^{n} \gamma_X \left[\mu\left(A \cup \{x_i\}\right) - \mu(A) \right], \tag{4}$$

where:

$$\gamma_X(A) = \frac{(|X| - |A| - 1)!\,|A|!}{|X|!} \tag{5}$$

The Shapley value with respect to the measure μ is a vector $\nu = [\nu_1, \nu_2, \ldots, \nu_n]$. It describes the global importance of every element by considering the effects of all subsets with and without the given element. According to the definition, the Shapley value has the property that the sum of all its components is 1, which can be formulated as $\sum_{i=1}^{n} \nu_i = 1$. Scaled by the factor *n*, the Shapley values greater than 1 indicate that the given element (criterion) is more important than the average.

The Shapley value ranges between 0 and 1. In essence, it measures how much a criterion contributes, on average, to all the coalitions of criteria.

Definition 5 (Interaction Index - $I_{i,j}$, [58]). Let μ be a fuzzy measure on X. The interaction index of the criteria x_i and x_j is defined by:

$$I_{i,j} = \sum_{K \subset X \setminus \{x_i, x_j\}} \frac{(|X| - |K| - 2)!\,|K|!}{(|X| - 1)!} \left[\mu\left(K \cup \{x_i, x_j\}\right) - \mu\left(K \cup \{x_i\}\right) - \mu\left(K \cup \{x_j\}\right) + \mu(K) \right] \tag{6}$$

The interaction index takes values from [-1, 1] interval, where negative (positive) values indicate a negative (positive, synergic) interaction.

Definition 6 (Ordered Weighted Mean - *OWA*, [88]). An OWA aggregation operator is a mapping *OWA*: $[0.1]^n \rightarrow [0.\ 1]$ such as:

$$OWA(x_1, \ldots, x_n) = \sum_{i=1}^{n} w_i x_{(i)}, \tag{7}$$

where the weights $w_i \in [0,1]$ for $i = 1, \ldots, n$, $\sum_{i=1}^{n} w_i = 1$ and $x_{(i)}$ indicates that the indices have been permuted so that $0 \le x_{(1)} \le \ldots \le x_{(n)} \le 1$.

3.3.3. Integrated assessment process for machine criticality identification

The operation process of the Choquet integral for the machine criticality criteria aggregation is described as follows (Figure 5).

The first step in developing a machine criticality index focuses on weighting the individual elements (criteria and sub-criteria). The assessment of the importance of criteria is usually a subjective assessment and is carried out by experts. The subjective approach requires evaluator(s) to evaluate the criteria in terms of a relative importance or influence of the criteria towards the final score [49]. Since this step is carried out by a team of experts and because of their subjectivity

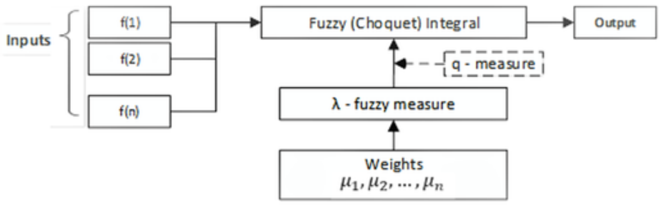


Fig. 5. λ-fuzzy measure and fuzzy integral

and cognitive differences, linguistic variables are used and then they are aggregated by fuzzy arithmetic.

The Fuzzy Number Ordered Weighted Average (FN-OWA) operator was used in the model for averaging expert evaluations. According to Sadiq and Tesfamariam [71]) the FN-OWA operator:

- provides a flexible aggregation ranging between the minimum and the maximum operators for fuzzy (or qualitative) data;
- has ability to aggregate not only the quantitative data but can also handle linguistic as well as crisp data;
- can handle the missing information efficiently, i.e., a case of complete ignorance about the value of a given input parameter;
- provides flexibility in handling *exaggeration* and *eclipsing* in the aggregation process;
- the aggregated value obtained through FN-OWA operator retains the same linguistic state as if all input criteria have equal values, i.e., idempotency property of the FN-OWA operator.

Aggregated fuzzy weights are then defuzzied in order to be applied in constructing a fuzzy measure. Mathematically, defuzzifying a fuzzy set is the process of rounding it off from its location to the nearest vertex, what reduces the set into the most typical or representative value. Compared with a fuzzy value, a crisp value is more intuitive and easier for the final comparison because fuzzy sets have partial ordering. These crisp values (fuzzy density) can be treated as an average assessment of the importance of individual criteria/sub-criteria.

The next step is to build a fuzzy measure λ. The fuzzy measure is an extension of a probability measure. Probability measures are usually resistant in representing human subjectivity because of their additivity. In contrast, fuzzy measures do not require this property and, thus, can be interpreted as the subjective measures of a person evaluating an object [81]. This kind of measure is more flexible than a probability. According to Beliakov et al. [3]: Fuzzy measures map each subset of a given set to a weight or importance, what allows for the modelling of complementary or redundant relationships between variables.

There are three kinds of interactions between the assessment criteria: synergy, inhibitory and non-interaction. The fuzzy measure can be applied to all three situations. In order to lower the number of coefficients (which increases exponentially with a number of criteria) and satisfy the monotonicity and continuity, the criticality machine assessment model uses λ-measure (see chapter …, Definition 2).

Let any subset $A_i = \{x_1, x_2, \ldots, x_i\}$ of X and given λ value (as calculated in Eq. (2)), the fuzzy measure $\mu_\lambda(A_n)$, for $1 \le i \le n$ can be determined recursively as:

$$\mu_\lambda(A_1) = \mu_\lambda(\{x_1\}) = \mu_1 \tag{8}$$

$$\mu_\lambda(A_i) = \mu_i + \mu_\lambda(A_{i-1}) + \lambda \mu_i \mu_\lambda(A_{i-1}) \tag{9}$$

In this application, the values of the fuzzy densities of the λ-measure are provided by experts according to their opinion on the worth of information sources. If experts choose to provide values that add to 1, the unique real value of the parameter λ will be zero, and, hence, the λ-measure will actually be a probability-measure even though this might not be the best measure for modeling the system.

One of the problems which can appear in case of expert assessments is the situation in which individual criteria will be rated so high (close to 1) that pairs, triples etc. of the criteria, due to the monotonicity of the fuzzy measure, will have very similar values (effectively equal to 1). While in case of aggregation using a weighted average such a situation is not a problem, in case of a fuzzy measure and the Choquet integral it can lead to undesirable results (shallowing / equalization of the criteria weights and total omission of interactions among criteria).

In order to reduce significantly the impact of these problems on the aggregation result, the q-measure proposed by Mohamed and Xiao [55] was applied. It is an extension of the Sugeno λ-measure that allows to automatically rescale the input density values μi. In practice, using a raw expert input is not a plausible strategy, because the values provided by experts, or obtained by using some computations, are at best on an interval scale with an arbitrary position of one. Therefore, scaling of these numbers is arbitrary, and computing from these numbers is then meaningless. The proposed definition for the q-measure, which is merely a normalization of λ -measure, solves this critical problem efficiently. The q-measure formulation decorrelates λ and the density. Moreover, such a formulation ensures that the q-measure complies with the principle that the fuzzy measure of any set, including the singleton sets, should not be determined by simply considering only that one set, regardless of the whole universe. This is a critical issue especially when we intend to find an appropriate fuzzy measure in order to model a complex system that manifests a high degree of interdependencies among its information sources.

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set. For all sets A, B ⊂ X with $A \cap B = \varnothing$, we define $\mu : 2^X \to [0,1]$ by:

$$\mu(X) = 1$$
$$\mu(A \cup B) = \mu(A) + \mu(B) + \lambda\mu(A)\mu(B) \tag{10}$$

for any choice of $\lambda \geq -1$. The only two constraints on the choice of a density generator value are:

$$0 \leq \mu_i \leq 1, i = 1,2,\ldots,n \tag{11}$$

$$\sum_{i=1}^{n} \mu_i > 0 \tag{12}$$

enforcing the density generators to have values in the unit interval with at least one of the values being strictly positive in order to insure a proper definition of the proposed fuzzy measure. Given a set of the density generator values $\{\mu_1, \mu_2, \ldots, \mu_n\}$ that satisfy the requirements (11) and (12) Mohamed and Xiao [55] defined the q-measure μ_q: $2^X \to [0, 1]$:

$$\mu_q(A) = \frac{\mu(A)}{\mu(X)} \forall A \tag{13}$$

It is called the q-measure because it is defined with the aforementioned quotient. Using Eq. (13), for any choice of the variable $\lambda \in [-1, \infty)$, Mohamed and Xiao [55] construct a fuzzy measure. This provides a definition for a class of various fuzzy measures specified by the choice of the variable. The Sugeno λ-measure is a special case in this class, when λ is selected such that μ(X) = 1. All fuzzy measures in this research are obtained by the application of this procedure to expert data.

It is important to note that in a fuzzy measure the importance of a single criterion or a pair of criteria is not solely determined by $\mu(\{x_i\})$ or $\mu(\{x_i, x_j\})$. One needs to consider all $\mu(A)$ such as $x_i \in A$ or $\{x_i, x_j\} \subseteq A$. Murofushi [59] and Murofushi and Soneda [58] proposed a solution to this problem based on the game theory for a single criterion and utility theory for pairs of criteria. Based on a fuzzy measure, the importance index (Shapley value) and interaction indices of different perspectives and criteria were defined.

When a fuzzy measure is constructed, the next step is to apply it in the Choquet integral to obtain the value of MCI. The Choquet integral (Definition 3) with respect to a fuzzy measure, compute an average of their inputs while also accounting for input interactions. This way, redundant inputs are not double counted while complementary inputs reinforce each other. Thanks to the stability of Choquet integral under positive linear transformations, the exact numerical scale in relation to which the calculations are made is not relevant. As such, the collection of the data from experts is a simplified way and allows for the assessment with the use of a linguistic scale.

4. Machine criticality assessment model for aviation industry – results of empirical studies

4.1. Development of the machine criticality assessment model

The research on the validity of the criteria for assessing the criticality of machines was carried out in the aviation industry. Based on the analysis of the research results (chapter 3.2), thirteen criteria were selected. They were considered important in this industry (the highest total value for the criteria and the greatest consistency of respondents' answers - standard deviation). These criteria are: "Machine replacement in case of a failure"; "Degree of influence on other tasks in case of a machine failure"; "Costs incurred by production for machine downtime (breakdown)"; "Number of nonconformities due to a machine failure during the year"; "Degree of machine influence on the final product quality"; "Costs of non-conforming products as the result of a failure"; "Frequency of failures per year"; "Average downtime of a technical facility due to failures and repairs"; "Spare parts availability"; "Failure elimination costs (excluding OHS and environmental costs)"; "In case of a failure, the degree of influence on the working environment"; "In case of a failure, the degree of risk to the environment"; "Machine failure costs due to health, safety and the environment".

A large number of criteria indicated by experts shows that the problem of a machine criticality assessment on a manufacturing system level is a complex multi-dimensional decision problem [79]In order to solve this problem, an enterprise has to consider different viewpoints from various stakeholders and, thus, include many (not always compatible) goals in a decision-making process. A possible strategy to deal with this problem is to combine multiple goals simultaneously into a hierarchical structure mapping the main stakeholder groups and the issues relevant to each of them. Considering the above, the problem of assessing machine criticality was structured as a hierarchy that shows the criteria and sub-criteria. This type of presentation is enterprise-friendly and enables more effective analysis.

The thirteen criteria are grouped into four categories: 1) Production - P, 2) Quality - Q, 3) Maintenance - M, and 4) Safety, Health and the Environment - SHE. The adopted categories of grouping (hereinafter referred to as criteria) reflect the main groups of stakeholders, i.e. those who are affected by the criticality of machines and which influence this criticality. The model was discussed with experts from the aviation industry and its final structure is shown in Figure 6.

Because of the formulated goal, the research was of qualitative nature. Qualitative research does not aim to draw a statistical inference or produce a statistically representative sample. Therefore, purposive sampling (also called judgment sampling) was used to select quality informants for this study by Tongco [83]. He asserted that there is no cap on how many informants should be considered in purposive sampling, but five is the minimum number for data to be reliable. According to Gray [24] and Guest et al. [25] a sample size of between six and twelve interviews is often sufficient to achieve data saturation for every theme. Experts from 8 aviation manufacturing companies were invited to participate in the research.

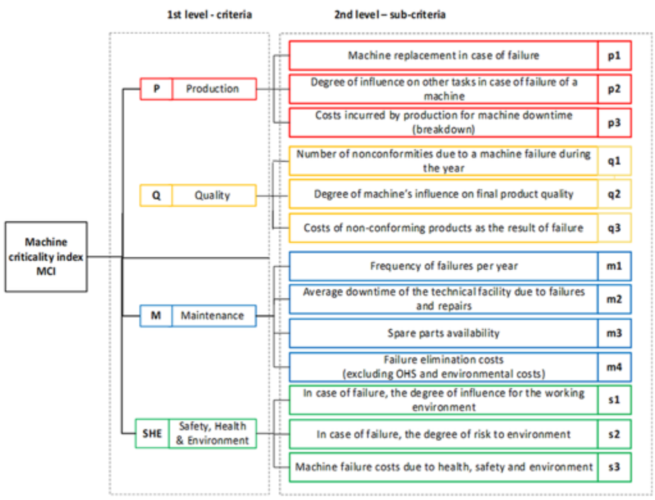


Fig. 6. Machine criticality assessment criteria and related sub-criteria

4.2. Determining the degree of importance of the machine criticality criteria and sub-criteria

The determination of the importance of criteria was carried out by experts according to the scheme (Figure 6):

(1) assessment of the importance of the criteria and sub-criteria by experts;
(2) calculating the an average value for the criteria and sub-criteria;
(3) developing a λ measure;
(4) determining the importance and interactions between the criteria and sub-criteria.

First, the experts evaluated the importance of the criteria/sub-criteria in a questionnaire. To increase the accuracy of the machine criticality assessment, all aerospace experts in our sample checked that the machine criticality frame used was working in their maintenance systems before conducting the assessment. The experts were asked to answer the following questions: How important is the X criterion / sub-criterion if it were to be used alone to assess the machine criticality? (following the hierarchical model - Figure 6, with the five-level linguistic scale (Table 3)). The experts had no imposed numerical interpretation of the linguistic variables used [91].

Table 3. Linguistic values of the criteria/sub-criteria importance grade

Linguistic terms	Description	Linguistic values
Very important	the criterion/sub-criterion can be used alone to assess the entire level	(0.75, 1.0, 1.0)
Important		(0.5, 0.75, 1.0)
Moderately important		(0.25, 0.5, 0.75)
Equal important		(0, 0.25, 0.5)
Irrelevant	the criterion/sub-criterion is almost irrelevant to the level assessment	(0, 0, 0.25)

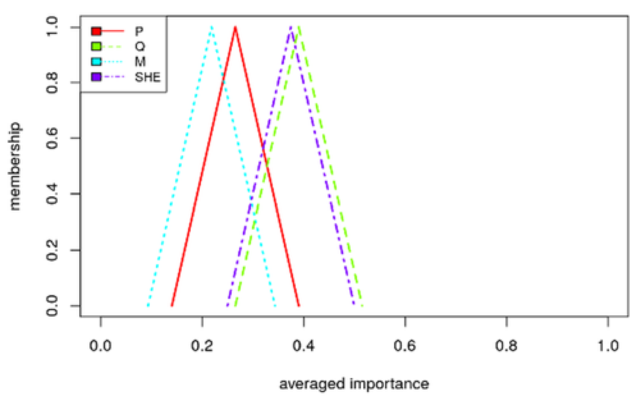


Fig. 7. Averaged importance values for the first level criteria

The average importance for each sub-criterion was calculated using FN-OWA. As a result of this aggregation method, the most extreme evaluations were rejected. The same method was used to aggregate expert assessments for the criteria (second level). Averaged importance values for the first level criteria are given in Figure 7.

Table 4. Fuzzy densities μi, fuzzy measure and Shapley value (ν_i) for sub-criteria

Criteria	Sub-criteria	$\mu_i(\bullet)$	λ	ν_i
Production	p1	0.3750	-0.0458	1.1085
	p2	0.3281		0.9690
	p3	0.3125		0.9225
Quality	q1	0.3125	0.1526	0.9840
	q2	0.3438		1.0800
	q3	0.2969		0.9360
Maintenance	m1	0.2250	0.1068	0.9360
	m2	0.2875		1.1920
	m3	0.2500		1.0388
	m4	0.2000		0.8332
SHE	s1	0.3594	0.0484	1.0944
	s2	0.3438		1.0473
	s3	0.2813		0.8583

The last step was to calculate single numerical values for each of the fuzzy numbers. These values (fuzzy densities μ_i) can be treated as an average assessment of the importance of individual criterion/sub-criterion. Using Center of Gravity defuzzyfication method led to the results presented in Table 4 and 5.

The calculated values of μ_i were used to develop a fuzzy measure. An algorithm presented in Mohamed and Xiao [55] was implemented in the R 3.4.4 Statistical Computing Platform and applied without fixing λ to the averaged importance values in order to construct the Sugeno λ-measure (Definition 2). Once the fuzzy measures for the sub-criteria and criteria are identified, the next step is to compute the Shapley value using Eq 4 and Eq 5. The obtained fuzzy densities μ_i, λ-values for the sub-criteria and criteria as well as the scaled Shapley value are presented in Table 3 and Table 4.

Table 5. Fuzzy densities μi, fuzzy measure and Shapley value (ν_i) for criteria

Criteria	$\mu_i(\bullet)$	λ	ν_i
Production	0.2656	-0.4788	0.8368
Quality	0.3906		1.2688
Maintenance	0.2188		0.6816
SHE	0.3750		1.2132

In Table 5, the λ value equals -0.4788212, what indicates a high degree of an interaction between various criteria for assessing the machine criticality. Based on the fuzzy measure (Table 4 and Table 5), the importance index (Shapley value) of different criteria and sub-criteria was defined.

The Shapley value measures a relative importance of each sub-criterion/criterion in terms of its contribution to the score of each coalition [7]. It can measure the importance of each feature in the contribution to the machine criticality assessing problem better. The results presented in Table 5 indicate that the most important criteria are ‘Quality’ (ν_i=1.2688) and “Safety, Health and the Environment” (ν_i=1.2132), whereas the ‘Maintenance’ criterion is the least important (ν_i=0.6816). The results presented in Table 3 apply to the value of the Shapley index for the sub-criteria describing a particular criterion. Analyzing the criterion ‘Production” (P), the experts indicate that “Machine replacement in case of a failure - p1” is more important than “Degree of influence on other tasks in case of a machine failure - p2” and “Costs incurred by production for machine downtime (breakdown) - p3”. Assessing the criterion “Quality” (Q) (Table 3), the experts indicate that the sub-criteria “Degree of machine’s influence on a final product quality - q2” is the most important. Another criterion analyzed is “Maintenance” (M). The distribution of the importance of the assessment sub-criteria indicates that the most important are “Average downtime of a technical facility due to failures and repairs - m2” and “Spare parts availability - m3”. The fourth criterion is ‘Innovation and development’ (ID). According to the experts’ assessment the most important criteria are “In case of a failure, the degree of influence on the working environment - s1” and “In case of a failure, the degree of risk to the environment - s2”

Another interesting aspect is that of the interaction among the criteria. When the fuzzy measure is not additive, then some criteria interact. The weight of sets of the sub-criteria taken together is determined by the Interaction Index, measuring the synergies or redundancies existing between the sets of variables. The obtained interaction index for the sub-criteria is presented in Table 6.

Table 6. Interaction Index $I_{i,j}$ for sub-criteria

Criteria	P	Sub-criteria	(p1, p2)	(p1, p3)	(p2, p3)
		$I_{i,j}$	-0.0056	-0.0053	-0.0047
	Q	Sub-criteria	(q1, q2)	(q1, q3)	(q2, q3)
		$I_{i,j}$	0.0168	0.0145	0.0159
	M	Sub-criteria	(m1, m2)	(m1, m3)	(m2, m3)
		$I_{i,j}$	0.0071	0.0062	0.0079
		Sub-criteria	(m1, m4)	(m2, m4)	(m3, m4)
		$I_{i,j}$	0.0049	0.0063	0.0055
	SHE	Sub-criteria	(s1, s2)	(s1, s3)	(s2, s3)
		$I_{i,j}$	0.0060	0.0049	0.0047

According to the assessment of the experts from the aviation industry, the sub-criteria describing the criterion “Production” are redundant, which means that some criteria should be rejected. Nevertheless, since the values of the interaction ratios are close to zero, it is difficult to draw binding conclusions. The interaction indexes between the sub-criteria describing the remaining criteria for assessing the criticality of machines are positive. Therefore, it can be assumed that they are synergistic (see chapter 3.3.2). The most complementary criteria are q1 and q2, and q2 and q3.

4.3. Numerical example

The above multicriteria criticality assessment model was applied to assess Machine Criticality Index (MCI) in a medium size aviation factory. The calculation procedure of MCI requires the fuzzy measure (μ) and actual values of the sub-criteria obtained from the company assessment team (f_i) (see an outline of this procedure in Fig. 5). Based on the available data collected in various departments of the company, the supervisor of the maintenance department assessed each of the 13 sub-criteria specified in the model for the selected machine A (Figure 6). To assess the value of the individual sub-criteria a point method was selected. The literature review [66, 75] indicates that this is the most common method of assessing criteria used in the aviation industry. Table 7 presents an assessment matrix for the example criteria.

Table 8 presents the value of the sub-criteria f_i. The aggregated values $C_\mu(f_1, f_2, \ldots, f_n)$ was obtained by Eq.(3) using the importance weighting of $\mu_i(\bullet)$ for the sub-criteria and for criteria (table 3) in R 3.4.4 Statistical Computing Platform. The aggregated value $C_\mu(f_1, f_2, \ldots, f_n)$ in Table 8 represents the overall criticality of the machine A of the four criteria: Production (P); Quality (Q), Maintenance (M) and (SHE). Based on Table 7, the Choquet integral values $C_\mu(f_1, f_2, \ldots, f_n)$ of each sub-criterion can further be employed to determine the next f_i and obtain the MCI for machine A (Table 9).

The output result is easy to interpret and understand, and can thus be used directly by all maintenance stakeholders. The value of $\mu_\lambda(\bullet)$ enables the assessment of the impact of each of the analyzed criteria on the final value of the MCI index for machine A. Among the analyzed criteria, „Quality” has the greatest impact (μ_λ (M, SHE, Q)−μ_λ (M, SHE) = 0.287). Therefore, in order to improve the maintenance strategies planned and implemented for machine A, first of all, actions should be defined in relation to the sub-criteria q1, q2 and q3. In this group of sub-criteria, q1 has the greatest impact (the highest value of ν_i=1.1085), therefore, potential solutions should be targeted at this area of impact.

5. Conclusions

Manufacturing equipment (machines, devices) are essential to production environments. However, due to the importance in the product realization process and the consequences of failure (e.g. environmental impact, human health and safety), not all machines are equally important. Given that each enterprise has limited resources (e.g. financial, human, material), it is necessary to prioritize (machine criticality assessment) and have a strategy to manage machines according to how critical they are to operation and maintenance.

In this paper the problem of machine criticality was analyzed. The criteria and methods proposed for assessing the criticality of machines were identified. Then, the research was conducted to identify the most important criteria used to assess the criticality of machines in various types of industries. On the basis of the obtained results, the criteria and industry, for which the machine criticality assessment model was developed, were selected.

The proposed model of the machine criticality assessment has a two – levelled hierarchical structure. On the first level of the hierarchical structure there are the criticality assessment process stakeholders. The criticality assessment process stakeholders are: maintenance managers who plan and realize maintenance activities and production, as well as quality and SHE managers, on whom the decisions and activities have impact. The second level of the hierarchical structure are sub-criteria – the aspects which are significant for all criticality assessment process stakeholders.

In order to assess the machine criticality the Machine Criticality Index was developed. The aim of the MCI index is to measure outputs of different criticality criteria and sub- criteria, and integrate them in one single index. Weighting and aggregation is an important step in this procedure. There are various weighting and aggregation methods related to specific purposes. Because the criteria and sub- criteria of the machine criticality assessment are independent, in order to aggregate them a non-additive fuzzy integral was selected. The fuzzy integral method applies fuzzy measures to deal with the problems of human subjective perception and uncertainty as well as to address the level of interdependency effects among the criteria [77]. In this research, we are motivated to implement the theory of fuzzy measures to model the

Table 7. The ranking of sub-criteria assessment – example

Ranking scale	Sub-criteria			
	p3	q2	s1	s2
	The failure			
1	has no effect on production losses at all	has no effect on product quality at all	has no effect on safety at all	has no effect on environment at all
2	can cause minor losses of production (p3 < a)	can create defects that will cause rejection or rework of parts of production lots	can cause only small injuries with no absence of the worker	can cause only a small impact only in the delimited area of occurrence inside the department
3	can cause significant losses of production a ≤ p3 < b	can create defects that will block online lots of production, causing high volumes of rejection or rework	can cause injuries with temporary absence of the worker	can cause an environmental impact internally in the plant
4	can cause extensive losses of production p3 ≥ b	can create defects that will be perceived by a customer (cannot be blocked inside the plant)	can cause death or injuries with permanent absence of the worker	can cause an environmental impact outside the limits of the plant

Table 8. The fuzzy measure and aggregated values of P, Q, M and SHE for the machine A

Criteria	Sub-criteria	f_i	$\mu_i(\bullet)$	$C_\mu(f_1, f_2, \ldots, f_n)$ (λ – value)	$\mu_\lambda(\bullet)$
Production	p1	1	0.3750	P = 1.636 λ = -0.0458	μ_λ(p2)=0.328
	p2	2	0.3281		μ_λ(p2, p3)=0.636
	p3	2	0.3125		μ_λ(p2,p3,p1)=1.000
Quality	q1	2	0.3125	Q = 2.000 λ = 0.1526	u_λ(q1)=0.313
	q2	2	0.3438		μ_λ(q1,q2)=0.673
	q3	2	0.2969		μ_λ(q1,q2,q3)=1.000
Maintenance	m1	3	0.2250	M = 2.519 λ = 0.1068	μ_λ(m1)=0.225
	m2	3	0.2875		μ_λ(m1,m2)=0.519
	m3	2	0.2500		μ_λ(m1,m2,m3)=0.783
	m4	2	0.2000		μ_λ(m1,m2,m3,m4)=1.000
SHE	s1	3	0.3594	SHE = 2.068 λ = 0.0484	μ_λ(s1)=0.359
	s2	2	0.3438		μ_λ(s1,s2)=0.709
	s3	1	0.2813		μ_λ(s1,s2,s3)=1.000

Table 9. Fuzzy measure and value of MCI for machine A

Criteria	f_i	$\mu_i(\bullet)$	$C_\mu(f_1, f_2, \ldots, f_n)$ **λ - value**	$\mu_\lambda(\bullet)$
Production	1.636	0.2656	MCI = 2.079 λ = -0.4788	μ_λ(M)=0.219
Quality	2.000	0.3906		μ_λ(M,SHE)=0.554
Maintenance	2.519	0.2188		μ_λ(M,SHE,Q)=0.841
SHE	2.068	0.3750		μ_λ(M,SHE,Q,P)=1.000

importance and interaction between the features in the Choquet integral. According to the best knowledge of the authors, there is a lack of such framework of the criticality machine assessment in the previous research. Based on the fuzzy measure, the importance index (Shapley value) and interaction index of different criteria and sub-criteria were defined. The analysis of Shapley values and interaction indexes demonstrate that the presented fuzzy machine criticality assessment is able to provide maintenance managers with a better understanding of the importance of individual criteria and sub-criteria in the assessment of the machine criticality and their impact on the final value of the MCI index. Taking into account the final value of the MCI index they are able to develop better planning of machine maintenance programmes and resources allocation.

The created model has some limitations. First of all, the model was developed only based on the research conducted in enterprises from the aviation industry. Secondly, in these enterprises only discrete manufacturing processes were realized. Therefore, some of the analzyed criteria cannot be significant for continuous manufacturing processes, e.g. the sub–criterion p1 (Machine replacement in case of a failure). Thirdly, the calculation of the MCI index from a mathematical point of view is complicated. Therefore, it could be a potential limitation of the application of this model in practice. Finally, the development of an intelligent manufacturing system and digital twin technology with rich sensor data and AI technique for diagnostics and prognostics would have a great influence on the calculation of the MCI index. Thus, carrying out relevant research is suggested to be continued.

References

1. Antosz K, Ratnayake RMC. Machinery Classification and Prioritization: Empirical Models and AHP Based Approach for Effective Preventive Maintenance. IEEE International Conference on Industrial Engineering and Engineering Management (IEEM) 2016; 1380-1386, https://doi.org/10.1109/IEEM.2016.7798104.
2. Azadeh A, Ghaderi SF, Ebrahimipour V. An integrated PCA DEA framework for assessment and ranking of manufacturing systems based on equipment performance. Engineering Computations 2007; 24(4): 347-372, https://doi.org/10.1108/02644400710748689.
3. Beliakov G, James S, Wu JZ. Value and Interaction Indices. Discrete Fuzzy Measures. Studies in Fuzziness and Soft Computing 2020; 382, https://doi.org/10.1007/978-3-030-15305-2_3.
4. Bengtsson M. Classification of machine equipment. International Conference on Maintenance Performance Measurement & Management, Luleå 2011.
5. Bevilacqua M, Braglia M. The analytic hierarchy process applied to maintenance strategy selection. Reliability Engineering & System Safety 2000; 70(1): 71-83, https://doi.org/10.1016/S0951-8320(00)00047-8.
6. Bokrantz J, Skoogh A, Berlin C, Wuest T, Stahre J. Smart Maintenance: An Empirically Grounded Conceptualization. International Journal of Production Economics 2020; 223:107534, https:/doi.org/10.1016/j.ijpe.2019.107534.
7. Buck A, Anderson D, Keller J, Wilkin T, Islam M. A Weighted Matrix Visualization for Fuzzy Measures and Integrals. IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), IEEE, Piscataway, NJ. 2020, https://doi.org/10.1109/FUZZ48607.2020.9177775.
8. Chen RQ, Ma SH. Production and Operations Management; Higher Education Press: Beijing, China, 1999; 302-304.
9. Chiou HK, Tzeng GH. Fuzzy multicriteria decision-making approach to analysis and evaluation of green engineering for industry. Environmental Management 2002; 30(6): 816-830, https://doi.org/10.1007/s00267-002-2673-z.
10. Choquet G. Theory of capacities. In Annales de l'institut Fourier 1954; 5: 131 - 295, https://doi.org/10.5802/aif.53.
11. Colapinto C, Jayaraman R, Abdelaziz FB, La Torre D. Environmental sustainability and multifaceted development: multi-criteria decision models with applications. Annals of Operations Research 2020; 293: 405-432, https://doi.org/10.1007/s10479-019-03403-y.
12. Costantino F, Di Gravio G, Tronci M. Integrating environmental assessment of failure modes in maintenance planning of production systems. Applied Mechanics and Materials 2013; 295: 651-660, https://doi.org/10.4028/www.scientific.net/AMM.295-298.651.
13. Daniewski K, Kosicka E, Mazurkiewicz D. Analysis of the correctness of determination of the effectiveness of maintenance service actions. Management and Production Engineering Review 2018; 9 (2): 20-25, https://doi.org/10.24425/119522.
14. Faisall M, Sharawi A. Prioritize Medical Equipment Replacement Using Analytical Hierarchy Process. IOSR Journal of Electrical and Electronics Engineering 2015; 10(3): 55-63.
15. Farinha Torres JM. Asset Maintenance Engineering Methodologies. CRC Press 2018, https://doi.org/10.1201/9781315232867.
16. Fernandez O, Labib AW, Walmsley R, Petty DJ. A decision support maintenance management system: Development and implementation. International Journal of Quality & Reliability Management 2003; 20(8): 965-979, https://doi.org/10.1108/02656710310493652.
17. Gass SI, Rapcsák T. Singular value decomposition in AHP. European Journal of Operational Research 2004; 154(3): 573-584, https://doi.org/10.1016/S0377-2217(02)00755-5.
18. Gluhak M. Equipment Reliability Process in Krško NPP. Journal of Energy 2016; 65(2): 32-40.
19. Gopalakrishnan M, Skoogh A. Machine criticality assessment for productivity improvement. Smart maintenance decision support. International Journal of Productivity and Performance Management 2019, 68(5): 858-878, https://doi.org/10.1108/IJPPM-03-2018-0091.
20. Gopalakrishnan M, Skoogh A. Machine criticality based maintenance prioritization: identifying productivity improvement potential, International Journal of Productivity and Performance Management 2018A; 67(4): 654-672, https://doi.org/10.1108/IJPPM-07-2017-0168.
21. Gopalakrishnan M, Subramaniyan M, Skoogh A. Data-driven machine criticality assessment - maintenance decision support for increased productivity, Production Planning & Control 2020, https://doi.org/10.1080/09537287.2020.1817601.
22. Grabisch M, Labreuche CA. Decade of Application of the Choquet and Sugeno Integrals in Multicriteria Decision Making. Annals of Operations Research 2010; 175: 247-286, https://doi.org/10.1007/s10479-009-0655-8.
23. Grabisch M. The Application of Fuzzy Integrals in Multicriteria Decision Making. European Journal of Operational Research 1996; 89: 445-456, https://doi.org/10.1016/0377-2217(95)00176-X.
24. Gray DE. Doing Research in the Real World. Sage Publications, London 2004.
25. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods 2006; 18(1): 59-82, https://doi.org/10.1177/1525822X05279903.
26. Guo L, Gao J, Yang J, Kang J. Criticality evaluation of petrochemical equipment based on fuzzy comprehensive evaluation and a BP neural network. Journal of Loss Prevention in the Process Industries 2009; 22(4): 469-476, https://doi.org/10.1016/j.jlp.2009.03.003.
27. Guo W, Jin J, Hu SJ. Allocation of Maintenance Resources in Mixed Model Assembly Systems. Journal of Manufacturing Systems 2013; 32 (3): 473-479, https://doi.org/10.1016/j.jmsy.2012.12.006.
28. Guo YZ, Dong J. The application of fuzzy clustering analysis in process equipment importance classification. International Journal of Pressure Vessels and Piping 1997; 71: 175-179, https://doi.org/10.1016/S0308-0161(96)00066-X.
29. Guofa Li, Yi Li, Zhang X, Hou C, He J, Xu B, Chen J. Development of a Preventive Maintenance Strategy for an Automatic Production Line Based on Group Maintenance Method. Applied Science 2018; 8: 1781, https://doi.org/10.3390/app8101781.
30. Gürbüz T, Alptekin SE, Alptekin GI. A hybrid MCDM methodology for ERP selection problem with interacting criteria. Decision Support Systems 2012; 54: 206-214, https://doi.org/10.1016/j.dss.2012.05.006.
31. Hijes F, Cartagena J. Maintenance strategy based on a multicriterion classification of equipments. Reliability Engineering & System Safety 2006; 91(4): 444-451, https://doi.org/10.1016/j.ress.2005.03.001.
32. Hu YC, Chen H. Choquet Integral-based hierarchical networks for evaluating customer service perceptions on fast food stores. Expert Systems with Applications 2010; 37: 7880-7887, https://doi.org/10.1016/j.eswa.2010.04.049.
33. Jaderi F, Saidi E, Anvaripour B, Nabhani N. Criticality analysis for assets priority setting of Abadan Oil Refinery using AHP and Delphi Techniques. International Journal of Engineering and Innovative Technology 2012; 2(6): 48-53.
34. Jagtap HP, Bewoor AK. Use of Analytic Hierarchy Process methodology for criticality Analysis of thermal Power plant equipment. in Materials Today: proceedings 2017; 4: 1927-1936, https://doi.org/10.1016/j.matpr.2017.02.038.
35. Jasiulewicz-Kaczmarek M, Żywica P. The concept of maintenance sustainability performance assessment by integrating balanced scorecard

with non-additive fuzzy integral. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 650-661, https://doi.org/10.17531/ein.2018.4.16.
36. Khanlari A, Mohammadi K, Sohrabi B. Prioritizing equipments for preventive maintenance (PM) activities using fuzzy rules. Computers and Industrial Engineering 2008; 54: 169-184, https://doi.org/10.1016/j.cie.2007.07.002.
37. Kłosowski G, Rymarczyk T, Kania K, Świć A, Cieplak T. Maintenance of industrial reactors supported by deep learning driven ultrasound tomography. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 138-147, https://doi.org/10.17531/ein.2020.1.16.
38. Konowrocki R, Chojnacki A. Analysis of rail vehicles' operational reliability in the aspect of safety against derailment based on various methods of determining the assessment criterion. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 73-85, http://dx.doi.org/10.17531/ein.2020.1.9.
39. Kozłowski E, Borucka A, Świderski A. Application of the logistic regression for determining transition probability matrix of operating states in the transport systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(2): 192-200, https://doi.org/10.17531/ein.2020.2.2.
40. Kozłowski E, Mazurkiewcz D, Kowalska B, Kowalski D. Application of multidimensional scaling method to identify the factors influencing on reliability of deep wells. In: Burduk A, Chlebus E, Nowakowski T, Tubis A. (eds) Intelligent Systems in Production Engineering and Maintenance. ISPEM 2018. Advances in Intelligent Systems and Computing 2019; 835: 56-65, https://doi.org/10.1007/978-3-319-97490-3_6.
41. Labib AW. A decision analysis model for maintenance policy selection using a CMMS. Journal of Quality in Maintenance Engineering 2004; 10(3): 191-202, https://doi.org/10.1108/13552510410553244.
42. Lewandowski J. Decision making processes in reliability and operation of continuous running technical objects. Technical University of Lodz Press, Lodz 2008.
43. Li G, Li Y, Zhang X, Hou C, He J, Xu B, Chen J. Development of a Preventive Maintenance Strategy for an Automatic Production Line Based on Group Maintenance Method. Applied Sciences 2018,8: 1751, https://doi.org/10.3390/app8101781.
44. Li L, Ni J. Short - Term Decision Support System for Maintenance Task Prioritization." International Journal of Production Economics 2009; 121 (1): 195-202, https://doi.org/10.1016/j.ijpe.2009.05.006.
45. Liu P, Liu J. Some q-rung orthopai fuzzy Bonferroni mean operators and their application to multi-attribute group decision making. International Journal of Intelligent Systems 2018; 33(2), 315-347, https://doi.org/10.1002/int.21933.
46. Lopes IS, Figueiredo MC, Sá V. Criticality evaluation to support maintenance management of manufacturing systems. International Journal of Industrial Engineering and Management 2020; 11(1): 3 -18, https://doi.org/10.24867/IJIEM-2020-1-248.
47. Loska A. Methodology of variant assessment of exploitation policy using numerical taxonomy tools. Management Systems in Production Engineering 2015; 2(18): 98-104, http://dx.doi.org/10.12914/MSPE-09-02-2015.
48. Lotovskyi E, Teixeira AP, Guedes Soares C. Availability analysis of an offshore oil and gas production system subjected to age-based preventive maintenance by Petri Nets. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (4): 627-637, http://dx.doi.org/10.17531/ein.2020.4.6.
49. Ma J, Fan ZP, Huang LH. A subjective and objective integrated approach to determine attribute weights. European Journal of Operational Research 1999; 112(2): 397-404, https://doi.org/10.1016/S0377-2217(98)00141-6.
50. Mahmood K, Shevtshenko E. Analysis of machine production processes by risk assessment approach. Journal of Machine Engineering 2015; 15(1): 112-125.
51. Mane AM, Dongale TD, Bapat MS. Application of Fuzzy Measure and Fuzzy Integral in Students Failure Decision Making. OSR Journal of Mathematics 2014; 10(6): 47-53, https://doi.org/10.9790/5728-10634753.
52. Marichal JL, Roubens M. Dependence between criteria and multiple criteria decision aid. 2nd International Workshop on Preferences and Decisions, Trento, Italy,TRENTO'98. 1998 (https://core.ac.uk/download/pdf/17190587.pdf).
53. Márquez AC, Herguedas AS. Learning about failure root causes through maintenance records analysis. Journal of Quality in Maintenance Engineering 2004; 10(4): 254 - 262, https://doi.org/10.1108/13552510410564873.
54. Márquez AC, León PM, Fernández JFG, Márquez CP, Campos ML. The maintenance management framework: A practical view to maintenance management. Journal of Quality in Maintenance Engineering 2009; 15(2):167-178, https://doi.org/10.1108/13552510910961110.
55. Mohamed MA, Xiao W. Q-measures: an efficient extension of the Sugeno lambda-measure. IEEE Transactions on Fuzzy Systems 2003; 11(3): 419-426, https://doi.org/10.1109/TFUZZ.2003.812701.
56. Mohebbi A, Achiche S, Baron L. A fuzzy-based framework to support multicriteria design of mechatronic systems. Journal of Computational Design and Engineering 2020; 7(6): 816-829, https://doi.org/10.1093/jcde/qwaa059.
57. Moncadatorres A, Leuenberger K, Gonzenbach R, Luft A, Gassert R. Activity classification based on inertial and barometric pressure sensors at different anatomical locations. Physiological Measurement 2014; 35: 1245-1263, https://doi.org/ 10.1088/0967-3334/35/7/1245.
58. Murofushi T, Soneda S. Techniques for reading fuzzy measures (III): interaction index. in: 9th Fuzzy System Symposium 1993: 693-696.
59. Murofushi T. Techniques for reading fuzzy measures (I): The Shapley value with respect to a fuzzy measure. in: 2nd Fuzzy Workshop, Nagaoka, Japan 1992: 39-48.
60. Ni J, Jin X. Decision Support Systems for Effective Maintenance Operations. CIRP Annals - Annals 2012; 61 (1): 411-414, https://doi.org/10.1016/j.cirp.2012.03.065.
61. Paszkowski W, Kotus J, Poremski T, Kostek B. Evaluation of sound quality features on environmental noise effects - a case study applied to road traffic noise. Metrology and Measurement Systems 2018; 25 (3): 517-531.
62. Pękala B, Dyczkowski K, Grzegorzewski P, Bentkowska U. Inclusion and similarity measures for interval-valued fuzzy sets based on aggregation and uncertainty assessment. Information Sciences 2020; 547:1182-1200, https://doi.org/10.1016/j.ins.2020.09.072.
63. Pertiwi S, Hermawan W, Prahmawati E. Maintenance Cost Reduction of Paddy Seed Production Machinery by Implementing Preventive Maintenance System.; IOP Conf. Series: Materials Science and Engineering 2019; 557: 012075, https://doi.org/ 10.1088/1757-899X/557/1/012075.
64. Pschierer-Barnfather P, Hughes D, Holmes S. Determination of asset criticality: a practical method for use in Risk-Based investment planning. 21st International Conference on Electricity Distribution Frankfurt, 6-9 June 2011 paper no 1013.
65. Qi HS, Alzaabi RN, Wood AS, Jani M. A fuzzy criticality assessment system of process equipment for optimised maintenance management. International Journal of Computer Integrated Manufacturing 2013; 28(1): 112-125, https://doi.org/10.1080/0951192X.2013.814160.
66. Ratnayake RMC, Antosz K. Development of a Risk Matrix and Extending the Risk-based Maintenance Analysis with Fuzzy Logic Procedia

Engineering 2017b; 182: 2- 610, https://doi.org/10.1016/j.proeng.2017.03.163.
67. Ratnayake RMC, Antosz K. Risk-Based Maintenance assessment in the manufacturing industry: minimisation of suboptimal prioritisation. Management and Production Engineering Review 2017a; 8(1): 38-45, https://doi.org/10.1515/mper-2017-0005.
68. Ratnayake RMC, Stadnicka D, Antosz K. Deriving an empirical model for machinery prioritization: mechanical systems maintenance. IEEE International Conference on Industrial Engineering and Engineering Management 2014; 1442-1447, https://doi.org/10.1109/IEEM.2013.6962649.
69. Roda I, Macchi M, Fumagalli L, Viveros P. A review of multi-criteria classification of spare parts: From literature analysis to industrial evidences. Journal of Manufacturing Technology Management 2014; 25(4): 528-549, https://doi.org/10.1108/JMTM-04-2013-0038.
70. Roy W. Priority management: new theory for operations management. International Journal of Operations and Production Management 1994; 14(6): 4-24, https://doi.org/10.1108/01443579410062059.
71. Sadiq R, Tesfamariam S. Developing environmental indices using fuzzy numbers ordered weighted averaging (FN-OWA) operators. Stochastic Environmental Research & Risk Assessment 2008; 22(1): 494-505, https://doi.org/10.1007/s00477-007-0151-0.
72. Saleh N, Sharawi A, Elwahed M, Petti A, Puppato D, Balestra G. Preventive maintenance prioritization index of medical equipment using quality function deployment. IEEE Journal of Biomedical and Health Informatics 2015; 19(3): 1029-1035, https://doi.org/10.1109/JBHI.2014.2337895.
73. Shahin A, Attarpour MR. Developing Decision Making Grid for Maintenance Policy Making Based on Estimated Range of Overall Equipment Effectiveness. Mathematical Models and Methods in Applied Sciences 2011; 5: 86, https://doi.org/10.5539/mas.v5n6p86.
74. Singh RK, Kulkarni M. Criticality analysis of power-plant equipments using the Analytic Hierarchy Process. International Journal of Industrial Engineering & Technology 2013; 3(4): 1-14.
75. Stadnicka D, Antosz K, Ratnayake RMC. Development of an empirical formula for machine classification: Prioritization of maintenance tasks. Safety Science 2014; 63: 34 - 41, https://doi.org/10.1016/j.ssci.2013.10.020.
76. Subramanian N, Ramanathan R. A review of applications of Analytic Hierarchy Process in operations management. International Journal of Production Economics 2012; 138(2): 215-241, https://doi.org/10.1016/j.ijpe.2012.03.036.
77. Sugeno M, Terano T. A model of learning based on fuzzy information. Kybernetes 1997; 6(2): 157-166, https://doi.org/10.1108/eb005448.
78. Sun L, Dong H, Hussain OK, Hussain FK, Liu AX. A framework of cloud service selection with criteria interactions. Future Generation Computer Systems 2019; 94: 749-764, https://doi.org/10.1016/j.future.2018.12.005.
79. Świderski A, Borucka A, Grzelak M, Gil L. Evaluation of the machinery readiness using semi-Markov processes. Applied Sciences 2020; 10(4), 1541: 1-15. https://doi.org/10.3390/app10041541.
80. Taghipour S, Banjevic D, Jardine A. Prioritization of medical equipment for maintenance decisions. Journal of the Operational Research Society 2010: 1-22.
81. Terano T, Asai K, Sugeno M. Fuzzy Systems Theory and Its Applications. Academic Press lnc. 1992: 137-145.
82. Teunter RH, Babai MZ, Syntetos AA. ABC classification: Service levels and inventory costs. Production and Operations Management 2010; 19: 343-352, https://doi.org/10.1111/j.1937-5956.2009.01098.x.
83. Tongco MDC. Purposive sampling as a tool for informant selection. Ethnobotany Research and Applications 2007; 5: 147-158, https://doi.org/10.17348/era.5.0.147-158.
84. Triantaphyllou E, Mann SH. Using the Analytic Hierarchy Process for Decision Making in Engineering Applications: Some Challenges. International Journal of Industrial Engineering: Theory, Applications and Practice 1995; 2(1): 35-44.
85. Velasquez M, Hester PT. An Analysis of Multi-Criteria Decision Making Methods. International Journal of Operational Research 2013; 10(2): 56-66.
86. Wang T, Mousseau V, Pedroni N, Zio E. An empirical classification-based framework for the safety criticality assessment of energy production systems, in presence of in-consistent data. Reliability Engineering and System Safety 2017; 157:139-151, https://doi.org/10.1016/j.ress.2016.08.021.
87. Wójtowicz A, Żywica P, Stachowiak A, Dyczkowski K. Solving the problem of incomplete data in medical diagnosis via interval modelling. Applied Soft Computing 2016; 47: 424-437, https://doi.org/10.1016/j.asoc.2016.05.029.
88. Yager RR. On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man, and Cybernetics 1988; 18(1): 183-190, https://doi.org/10.1109/21.87068.
89. Yan S, Ma B, Wang X, Chen J, Zheng C. Maintenance policy for oil-lubricated systems with oil analysis data. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (3): 455-464, http://dx.doi.org/10.17531/ein.2020.3.8.
90. Yinghua HYL. Targeted Preventive Maintenance of Pharmaceutical Equipment. Journal of Drug Design and Medicinal Chemistry 2018; 4(2): 10-15, https://doi.org/10.11648/j.jddmc.20180402.11.
91. Zadeh LA. Fuzzy sets. Information and Control 1965; 8(3): 338-353, https://doi.org/10.1016/S0019-9958(65)90241-X.
92. Zuber N, Bajrić R. Gearbox faults feature selection and severity classification using machine learning. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (4): 748-756, http://doi.org/10.17531/ein.2020.4.19.

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl
Article citation info:
Sawczuk W, Merkisz-Guranowska A, Rilo Cañás A-M. Assessment of disc brake vibration in rail vehicle operation on the basis of brake stand. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 221–230, http://doi.org/10.17531/ein.2021.2.2.

Assessment of disc brake vibration in rail vehicle operation on the basis of brake stand

Indexed by: Web of Science Group

Wojciech Sawczuk[a], Agnieszka Merkisz-Guranowska[a], Armando-Miguel Rilo Cañás[b]

[a] Poznan University of Technology, Institute of Transport, ul. Piotrowo 3, 60-965 Poznan, Poland
[b] DB Systemtechnik GmbH, Adlergestell 143, 12439 Berlin, Germany

Highlights

- In the operation of a disc brake, the geometry of the system changes with the wear of the pads.
- The lever system is characterized by an uneven distribution of masses on the right and left side of the disc.
- There is a difference in the vibroacoustic signals of the right and left brake pads.
- The heavier side of the system shows greater friction in the joints and a reduction in pad vibration.
- The non-uniform vibrations of the friction pads indicate a disturbance in the braking process.

Abstract

The scientific aim of the article is to present the relationship between the vibroacoustic signals of the right and left friction pad during braking, depending on the mass distribution, as an element of the lever system. This article presents the results of tests of a railway disc brake in the scope of vibrations generated by pads in various states of wear located on both sides of the brake disc. The tests were carried out on the brake stand using the vibroacoustic method including the analysis of amplitudes and frequencies and the thermal imaging method. Special attention was paid to the analysis of the classic lever mechanism as a multi-mass system influencing the thermo-mechanical characteristics and vibrations of the pads on the right and left side of the brake disc. Uneven mass distribution of the system translates into uneven wear of the friction components. The scientific aim of this paper is to present the relation between vibroacoustic signals of the right and left friction pad during braking depending on the mass distribution of the lever system component.

Keywords

railway disc brake, brake vibration, lever mechanism.

1. Introduction

When scrutinizing the available literature sources, one may observe that the vibroacoustic processes generated by brake systems are analyzed in three directions. In the first, most numerous part, are works treating on the noise accompanying the braking process. In this respect, some of the researchers attempt to explore and identify the reasons for the occurring noise [8, 14, 15, 28, 41, 43]. Other researchers attempt to model the noise depending on the geometry features of the brake system and the components of the friction pair. The second, less numerous group of papers treating on the vibroacoustic processes in brake systems, constitute works on application of brake system vibrations in diagnostics of the wear level of a friction pair [29, 30]. The least numerous group of papers treating on vibroacoustics are works describing the use of brake system vibrations to evaluate the braking process itself. Simple analyses in the domain of amplitudes and frequencies to build regressive diagnostic models was applied in [28] that, upon transformation, allow assessing the value of the average coefficient of friction for selected braking speeds. Majority of models describing the vibrations of a brake system is based on the assumption that the increment of vibrations (a phenomenon heavily depending on a multitude of variables) is most dependent on the variability of the coefficient of friction between the brake pad and the brake disc. Additionally, the sensitivity of the brake system components (brake pads in particular) increases or reduces the vibrations generated by the brake system and its emission to the environment, which is described in [19] as a strongly unfavorable phenomenon, and to a decrease in vibrations generated by the braking system, which is a favorable phenomenon. The collected results from the operation of the braking system enable the analysis of its reliability as described in [1, 9, 16, 32], and, consequently, the development of algorithms for estimation of the time to failure as presented in [31].

The vibroacoustic processes generated by friction in brakes refer to simple models in the literature. These are two-mass systems with only the friction pad and the disc considered. However, there is no description of the vibroacoustic phenomena generated by friction materials in complex braking systems in automotive or railway vehicles. In such cases, apart from the model of contact between the friction pad and the brake disc itself, the lever system of force transmission from the brake cylinder to the friction pads should also be taken into account.

2. Vibration models and a model of the disc brake lever system

The initial models assumed that self-induced vibration of a brake was related to the drop in the coefficient of friction and increased slip velocity [14, 22, 24]. These were models of the elastic friction system called the stick-slip phenomenon. On this basis, it was observed that

E-mail addresses: W. Sawczuk - wojciech.sawczuk@put.poznan.pl, A. Merkisz-Guranowska - agnieszka.merkisz-guranowska@put.poznan.pl
A-M. Rilo Cañás - Armando-Miguel.Rilo-Canas@deutschebahn.com

the necessary condition for vibroacoustic phenomena to take place in a brake is the dependence of the coefficient of kinetic friction on the increase in the velocity $\mu_o(v_2)$ assuming inequality (1):

$$\frac{d\mu_o}{dv_2} < 0 \tag{1}$$

where: μ_o – coefficient of kinetic friction,
v_2 – velocity.

According to [14], if $\mu_2 > c/mg$, a self-induced vibration takes place generating the vibroacoustic phenomena, which, in the case of friction brakes, results in their squealing. It should be emphasized that the model of the stick-slip phenomenon of vibrations in the brake is correct in the case of a 1Bg single-block block brake. However, it does not work for the 2Bg double-unit block brake and disc brake.

Fosberry and Holubecki in [7] published theses that vibration in brake systems is caused by the cooperation of the brake pad-brake disc friction pair of the coefficient of static friction μ_{st} greater than the coefficient of kinetic friction μ_o or when the coefficient of kinetic friction increase with the increase in velocity v2. Similar conclusions were stated by Sinclair [34] and by Earles [4]. Other researchers such as Mills, Bowden and Leben [2] conducted research on resilient friction systems comparing them to the stick-slip motion [43, 44]. Eventually, the researchers declared that the vibroacoustic phenomena were not fully explored but their most likely explanation is the stick-slip motion in the friction coupling, whose source of energy is the change in the coefficient of friction as a function of velocity.

Spurr, in [37] proposes a term 'sprag-slip' to describe vibroacoustic phenomena in brake systems. He claims that the vibroacoustic phenomena generated by brake systems result from the contact of the friction pad with the brake disc. The friction force in such a system may be much greater than the same force in an ideally rigid system. In an actual brake system, due to elastic deformations and displacements, cyclic instantaneous drop and increase in friction occurs. This particular case was described by Spurr as spragging. Later, this model was improved by Jarvis and Earles [4, 14].

This was the first attempt of a theoretical explanation of the sprag-slip phenomenon, in which, in order to explain vibroacoustic phenomena, a rotating plate with a support was used. Later models, based on the Jarvis, Mills and Earles considerations, were more complex, had more degrees of freedom and several models of friction [21]. It should be emphasized that in the case of the sprag-slip phenomenon, despite the unilateral impact of the brake block (slider) on the brake disc, the researchers introduced a variable in the form of the slider inclination angle. It is a model of vibroacoustic phenomena the most similar to the real braking system. The angle of inclination of the slider corresponding to a change in the setting of the friction pads in relation to the brake disc has been taken into account in the model [36].

As Crolla and Lang [3, 17], have proven, this and other models do not entirely reproduce the actual brake. Yet, thanks to these models, one may obtain a qualitative indication for the process of design and search for solutions eliminating some classes of brake vibration, hence, the generated noise.

Then, Lang and Smales in [17] distinguished two types of vibroacoustic incidents originating in the brake systems. These distinctions are applied to date (phenomena occurring at low frequencies i.e. from 1 to 5 kHz and phenomena occurring at high frequencies, above 5 kHz.) The Lang's and Smale's model for low frequencies admits brake pads as rigid bodies. At high frequencies, one needs to additionally allow for deformations of the friction components. The assumption of a rigid body as a friction material was also utilized by Brooks [3], Milner [20] and Rudolph and Popp in their works [25, 26].

It should be emphasized that the models of vibroacoustic phenomena in brakes described in the literature concern a single disc and single pad system. In reality, a friction disc brake consists of a rotating disc to which friction pads are pressed from both sides. The models available in the literature analyse only the case of the friction pad acting on one side of the disc, assuming that the same phenomena and relations will occur on the other side of the brake disc. In this paper, an attempt has been made to present vibrations generated by friction pads on two sides of a brake disc, which differ from each other in amplitude values, and to present an analysis of probable reasons for the difference in these vibrations. Additionally, results of friction-mechanical and thermovision tests with significant differences in temperature values of individual friction linings are presented.

Railway disc brake is composed of a brake disc fixed to the axle and a lever mechanism [33]. The lever mechanism is composed of two main levers (right and left) connected with a central lever in the middle. On one end of the main lever, brake pad holders with the brake pads are fitted as well as the cantilevers and on the other end of the main lever there is a brake piston rod. The lever mechanism is fixed to the bogie frame at three points i.e. through the central lever and two cantilevers. Figure 1a and 1b presents the lever mechanism of a railway disc brake in two planes (side and top), while Figure 2 shows the diagram of the lever mechanism as a multi-mass model in the x-y and z-x plane.

When analyzing Figure 1c, it is noteworthy that, as the wear of the brake pads increases, in order to keep the constant distance of the brake pad and the disc (approx. 1mm [35]) when the brake is not active, the angles of the right and left levers must vary in the range α_1=0-13°. This is an interval of the angle value for the maximum admissible brake pad wear to the level of 5 mm (thickness of a new brake pad is 35 mm). Changes in the angle α_1 directly influence the values of forces acting on the brake pad holders.

The operation of the lever system in the z-x plane was then assessed. A diagram of the lever mechanism as a multi-mass model in this plane is shown in Figure 2b.

In the z-x plane, aside from the changes of the value of angle α_2 (Fig. 3b and 3c) due to the wear of the brake pads, additionally the action of the friction force on the contact point of the brake pads and the brake disc generates an inertia force F_{IW}, which makes part of the lever mechanism drop or lift depending on the direction of the disc rotation. As a consequence, on the vertical levers (cantilevers) clamping or separating forces are applied. If the lever mechanism is lifted when the friction force is directed upwards, the upper brake pads in

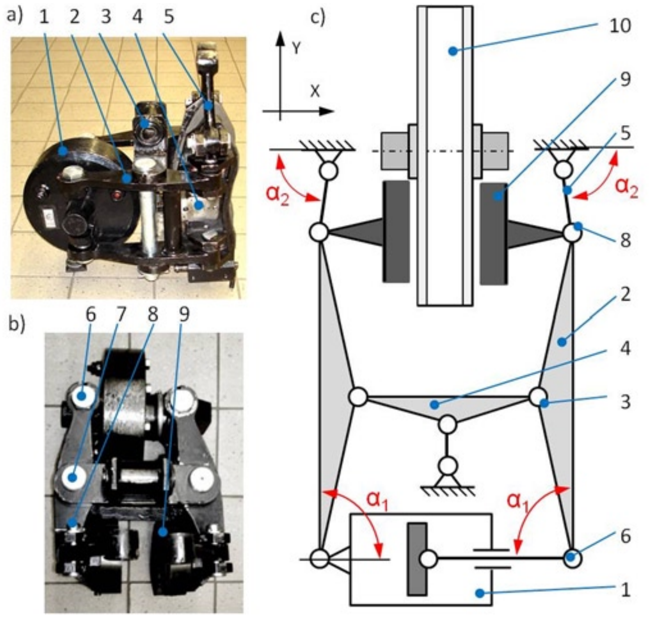


Fig. 1. View of the lever disc brake: a) side view, b) top view, 1-brake cylinder, 2-main lever (right and left), 3-central lever, 4-brake pad holder, vertical lever (cantilever), 6-brake cylinder pin, 7-central pin, 8-lever pin, 9-brake pad, 10-brake disc, c) structural model of the disc brake system

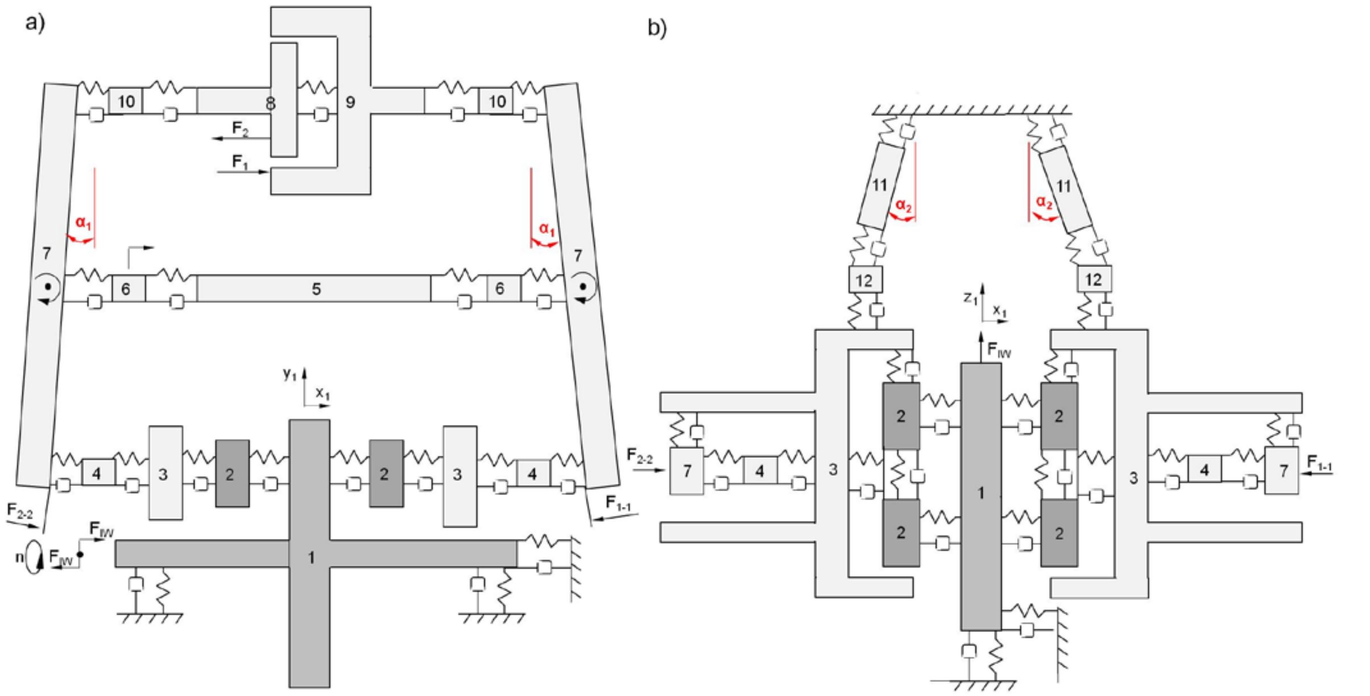


Fig. 2. View of the lever mechanism with visible parts as a multi-mass model in the plane: a) x-y b) z-x, 1- brake disc, 2- brake pad, 3- brake pad holder, 4, 6, 10, 12- pin, 5- central lever, 7- main lever (left, right), 8- piston rod, 9- cylinder, 11- cantilever, P_c- pressure in the cylinder, F_1- force from the piston rod, F_2- force from the cylinder, $F_{1-1,\ 2-2}$- forces acting on the brake pad holder with the brake pads, F_{IW}- inertia force on the wheel circumference during braking, n- disc rotational speed, α_1- angle of rotation of the main lever, α_2- angle of rotation of the cantilever

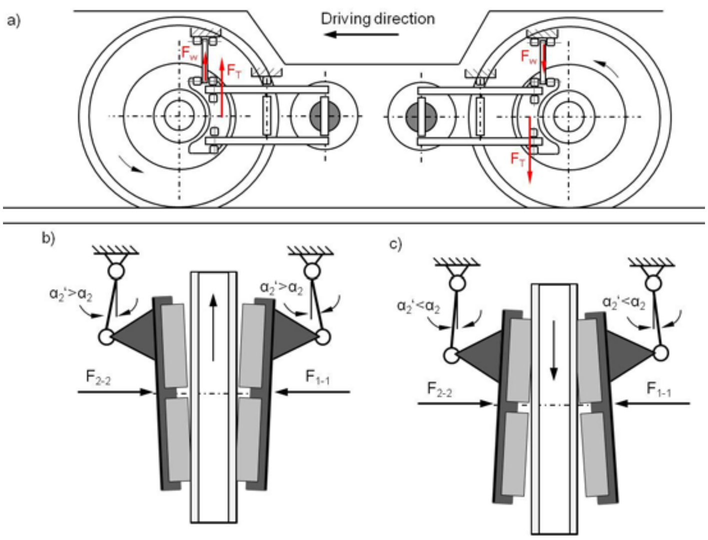


Fig. 3. Schematics of the bogie with a lever mechanism, a) forces acting on the lever mechanism during braking for a given driving direction, b) schematics of the brake pad setting when the friction force acts upwards, c) schematics of the brake pad setting when the friction force acts downwards, α2'-angle of the vertical lever during braking, α2- angle of the vertical lever at the moment of stopping v=0

the holder separate from the disc which increases the inclination angle α_2 of the vertical levers When the friction force is directed downwards, a separating force acts on the vertical levers and the lower brake pads in the holder are separated from the brake disc and the angle α_2 decreases. Figure 3 presents a diagram of the forces in the lever mechanism depending on the direction of rotation.

When analyzing the models presented in Figures 2a and 2b, one can observe that the railway disc brake, despite its simple design, is a multi-mass system [27]. Earlier research of the authors in the scope of identification of brake system vibration under laboratory (vibration frequency 4.5 kHz) and actual operation on a passenger railcar (vibration frequency 6 kHz) as well as those presented in [28, 30], confirm that vibroacoustic phenomena occur at high frequencies. According to the Lang's and Smales's theory, friction components are masses susceptible to deformation. Other researchers such as Rudolph, Popp and North [23, 26] built a model of a disc brake caliper of a passenger vehicle of two degrees of freedom. Later, they expanded the model to six components, also of two degrees of freedom. They have observed that the vibroacoustic phenomena are influenced by the complexity of the brake system – the more components in the brake lever mechanism, the more components potentially generating vibroacoustic phenomena [26]. Besides, in [30] the authors draw attention to the fact that the main components of a brake system have a common resonance frequency, which also may contribute to the generation and amplification of vibroacoustic phenomena.

4. Analysis of the results of measurements of the masses of individual components of the disc brake lever mechanism

In order to assess the mass distribution of the complete brake caliper, all brake system components were dismantled and then weight. Table 1 presents the masses of individual parts of the lever mechanism divided into left and right side of the clamping mechanism.

When analyzing Table 1, it can be observed that the sums of masses of individual components of the lever mechanism on the right and left side are not the same. The uneven distribution of the masses in the lever mechanism is influenced by the mass of the main levers (left and right) and the mass of the piston, in which the mass of the piston rod is almost twice as high as the cylinder. This is partly related to the additional mechanism of the piston rod, whose task is to set a constant clearance between the pads and the brake disc after braking (brake is disengaged) irrespective of the brake pad wear. Then, in the process of operation of the railway disc brake, as the brake pads wear, the angle α_1 increases on the main levers of the mechanism from 0 to 13° (in the case of the mechanism under analysis) and α_2.

Table 1. Masses in kilograms of individual components of the clamping mechanism

No.	Component		Symbol	Mass on the left side	Mass on the right side
1	Brake disc		m_1	116	
2	Brake pad	Upper	$m_{2,U,L}$, $m_{2,U,P}$	1.760	1.755
		Lower	$m_{2,D,L}$, $m_{2,D,P}$	1.745	1.750
3	Brake pad holder (caliper)		$m_{3,L}$, $m_{3,P}$	4.580	4.534
4	Brake pad holder pin	Upper	$m_{4,U,L}$, $m_{4,U,P}$	0.858	1.452
		Lower	$m_{4,D,L}$, $m_{4,D,P}$	0.294	
5	Central lever		m_5	18.540	
6	Central lever pin		$m_{6,L}$, $m_{6,P}$	1.238	1.325
7	Main lever		$m_{7,L}$, $m_{7,P}$	6.963	7.205
8	Pneumatic brake piston	Piston rod	m_8	**19.995**	-
9		Cylinder	m_9	-	**7.255**
9'	Complete brake piston		m_8+m_9	27.25	
10	Pin of the brake piston	from cylinder (top and left)	$m_{10,U,L}$, $m_{10,D,L}$	0.259 i 0.258	-
		from the piston rod	$m_{10,P}$	-	0.275
11	Vertical lever (cantilever)		$m_{11,L}$, $m_{11,P}$	5.237	5.127
12	Cantilever pin		$m_{12,L}$, $m_{12,P}$	0.370	0.372
13	Washers, safety pins		m_x	0.182	0.088
14	Sum of components on the left/right side		$m_{L,R}$	43.739	31.138
15	Difference in left and right weight		r_w	12.60	
16	Complete lever mechanism		$m_{2\text{-}13}$	93.42	

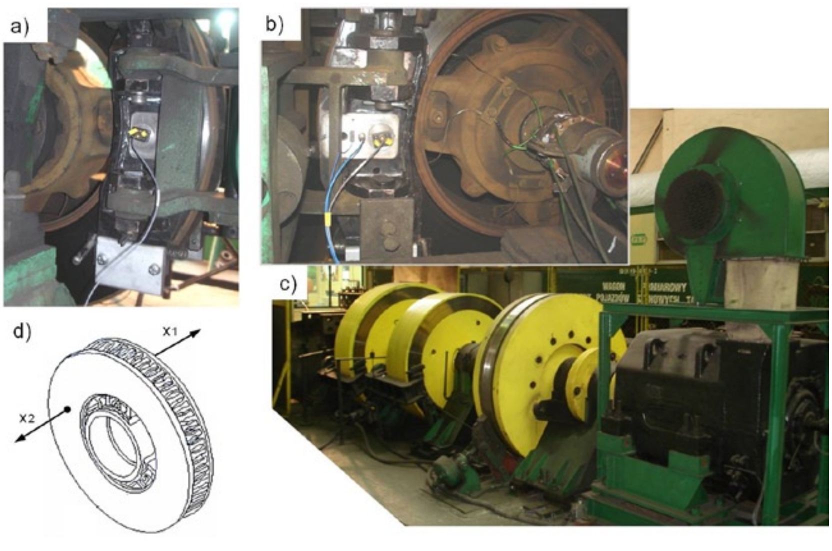


Fig. 4. Object of tests at the brake stand for testing railway disc brakes: a) view of the left brake support with a vibration transducer, b) view of the right brake support with a vibration transducer, c) view of the driving part of the brake position with rotating masses, d) orientation of vibration measurement directions

5. Methodology of tribological and vibroacoustic research

Friction-vibroacoustic research was conducted on the basis of the assumptions of the active experiment. During the tests, the input parameters, i.e. the state of the braking system, understood as wear of the friction pads, as well as such parameters as the braking start speed, pads pressure on the disc, mass to brake, were changed intentionally and in a specific way. Then, their impact on changes in output parameters was observed, such as the instantaneous and average friction coefficient, and acceleration of friction pad vibrations from the right and left of the brake disc.

The tests were carried out at a certified inertia brake stand, located at the Siec Badawcza Lukasiewicz - TABOR Rail Vehicles Institute in Poznan (Fig. 4). It is possible to perform tests on the rail block brake and disc brake reflecting the real conditions that occur when braking a rail vehicle. In addition, the Flir e60 thermal imaging camera was used during the tests to observe the temperature distribution of the friction linings after stopping braking.

The tests covered a ventilated brake disc with dimensions of Ø610×110 made of gray cast iron. The brake disc has been prepared for tests in accordance with the standard [5, 6]. In accordance with the manufacturer's procedure and the requirements contained in the code [42], the pads were made of thermosetting resin, synthetic elastomer, metal and organic fiber as well as friction modifiers.

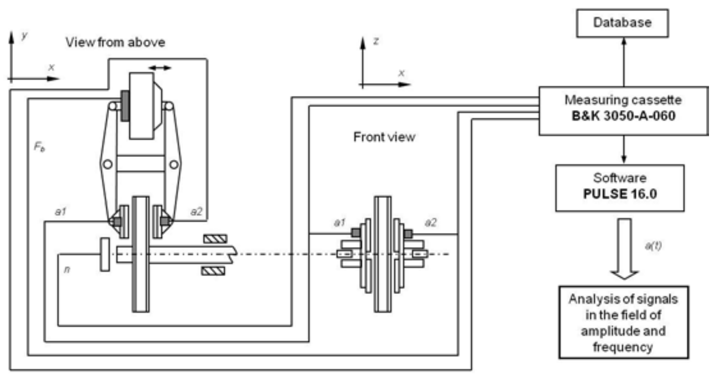


Fig. 5. Diagram of the measuring track used during vibroacoustic tests

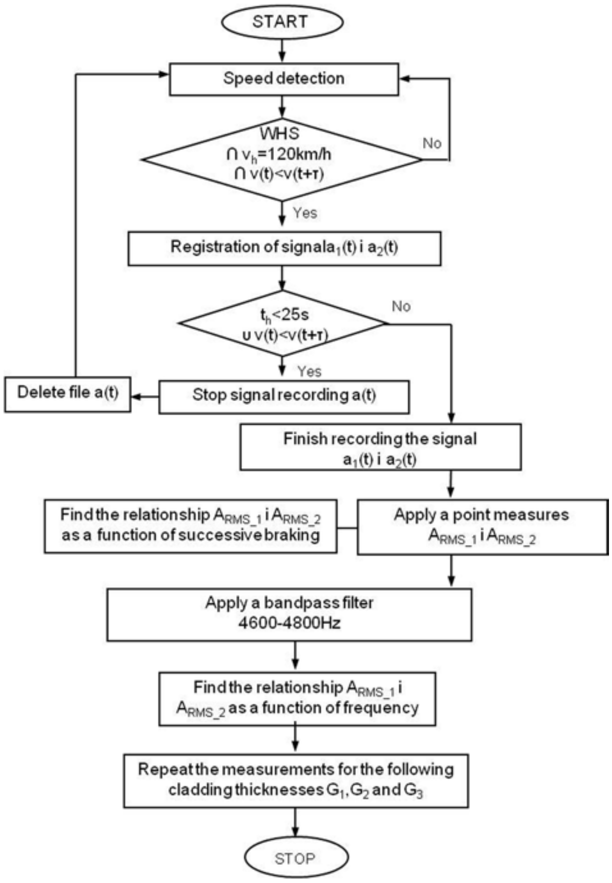


Fig. 6. Algorithm of vibroacoustic brake stand, WHS – perform stopping braking, τ – time gain

Three sets of pads were used for bench tests. The first new set of pads (4 pieces) with a thickness G_1 = 35 mm and two sets worn down to a thickness of G_2 = 25 mm and G_3 = 15 mm. Friction pad masses were m_{G1} = 1.75 kg (new pad), m_{G2} = 1.45 kg (pad worn up to 25 mm thick), m_{G3} = 1.02 kg (pad worn up to 15 mm thick), respectively.

Vibroacoustic tests were carried out in parallel with friction (tribological) tests. One vibration transducer (Fig. 4a and b) was attached to the brake mountings (right and left). The study was also carried out in accordance with the assumptions of the active experiment, where the input parameters were intentionally changed to record the output signals. The input values were simulated braking start speed v_o, brake pad pressure N, brake mass M, friction pad thickness G, and the output signals were instantaneous value vibration acceleration a. Then it was possible to observe the impact of input parameters on changes in output signals. B&K 4504A transducers were used to measure the vibrations. Figure 5 shows a diagram of the measuring track at the brake stand, additionally extended with the measurement of vibration acceleration.

When choosing the measurement site, the principle was adopted that the transducers should be located closest to the place generating the vibroacoustic signal associated with the operation of the brake friction pair and in a place easily accessible for measurement. The acceleration of vibrations was measured in a direction perpendicular to the surface of the brake disc, based on the experience of other researchers presented in [28, 38]. Figure 6 shows the algorithm of vibroacoustic brake stand.

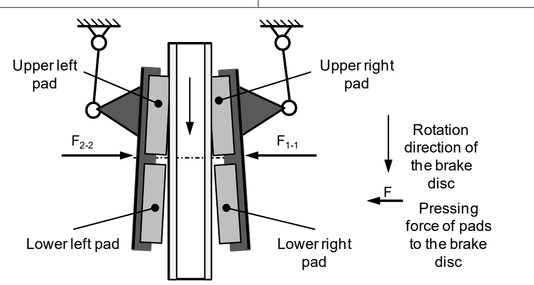
6. Analysis of tribological test results

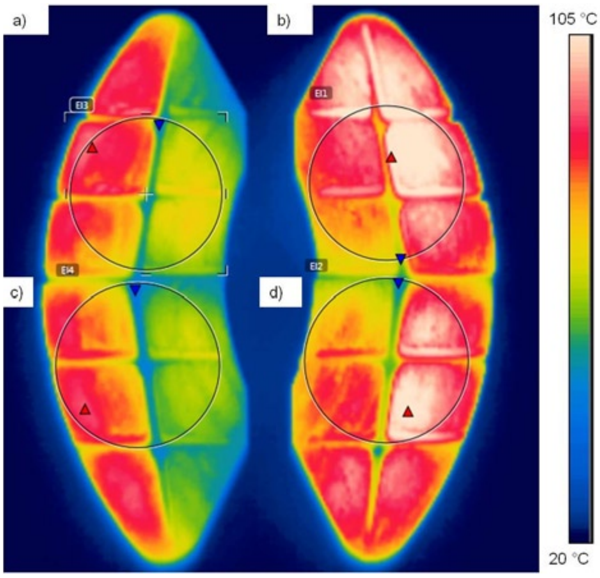
Friction tests carried out at a certified brake stand during braking in various combinations of brake lining pressure, braking masses, speed and degree of wear of the friction linings proved that in each braking case, wear is uneven on the piston rod and brake cylinder side. Also on the same side of the brake disc, the wear of the upper linings (above the rotational axis of the disc) as well as the lower ones varies. The set of weight consumption after 40 brakings for the lining pressure to the disc N = 44 kN, braking mass M = 7.5 t for five braking start speeds (50, 80, 120, 160 and 200 km/h) repeated 8 times, shown in Table 2. Wear tests of the friction material also included less lining pressure on the disc (28 kN) and less mass to be braked by one disc, i.e. 4.4 t. Similar studies on motor vehicles are described in [39].

In addition, during braking at the brake stand, there was a case of work of the stand as in the diagram shown in Figure 3c, i.e. the rotation of the brake disk pulls down the brake caliper. During the tests it was found that in each case of braking, both right and left bottom pads located in the brake mount wear slower. It was observed that in each case of braking, pads from the piston rod side (right side of the brake caliper) showed greater wear than pads from the brake cylinder side (left side of the caliper).

Due to the difficulty in determining the coefficient of slip friction, in the laboratory tests of the instantaneous coefficient of friction of the left and right side of the disc, the temperature distribution on the brake pads was determined with the thermo-visual methods. Figure 10 presents the temperature distribution on four brake pads after 40 instances of braking with the pressure N=44 kN and M=7.5 t.

Table 2. Consumption in grams of friction pads after 40 brakes with a pressure of N = 44 kN and a mass to brake M = 7.5 t

Weight consumption in grams of friction pads			
The pads New, thich G_1=35 mm		Pads worn to thickness G_2=25 mm	
Left side of the disc (from the brake piston rod)	Right side of the disc (from the brake cylinder)	Left side of the disc (from the brake piston rod)	Right side of the disc (from the brake cylinder)
Top pad	Top pad	Top pad	Top pad
117	129	115	123
Lower pad	Lower pad	Lower pad	Lower pad
112	114	105	112
Pads worn to thickness G_3=15 mm			
Left side of the disc (from the brake piston rod)	Right side of the disc (from the brake cylinder)		
Top pad	Top pad		
97	119		
Lower pad	Lower pad		
93	105		



Pads temperature °C			
From the piston rod side		From the cylinder side	
a)	77.0	b)	92.8
c)	73.9	d)	87.2

Fig. 7. Temperature distribution on the brake pads, a) upper left pad, b) upper right pad, c) lower left pad, d) lower right pad

When analyzing Figure 7, it is observed that the brake pads on the side of the piston rod have a higher temperature compared with the pads on the side of the cylinder. Besides, it can be observed that the lighter part of the brake cylinder i.e. piston rod, allows a better contact of the brake pads with the brake disc than the heavier brake cylinder. Higher temperature of the brake pads on the side of the piston rod translate into greater wear also on that side compared with the brake pads on the side of the cylinder. In addition, it was found that the difference in weight (12.6 kg) between the two brake cylinder assemblies also affects the frictional resistance in the bolt connections. It was observed that the right brake lever with the brake cylinder slidingly mounted on the central lever shows less resistance to movement in relation to the left brake lever with the piston rod.

7. Analysis of the result of vibroacoustic investigations

In the domain of amplitudes of analysis of vibration accelerations, point measures are most frequently applied that, with a single value, characterize a given vibration process in compliance with [29, 38]. Therefore, particularly in the vibroacoustic diagnostics (VD) it is possible to determine changes in the VD signal resulting from the change of the technical state of an object. There is a variety of papers available in literature presenting the application of vibroacoustic diagnostics in vehicles such as passenger vehicles, railway vehicles or aircraft [11, 12, 13, 22, 40]. In order to determine the relation between the average coefficient of friction and the vibration generated by the brake system, in the first place the authors have confirmed that there is a relation between the vibration measured on the brake pad holders and the condition of the brake system understood as the brake pad wear. To this end, on the test stand, for all velocities under analysis, instantaneous vibration accelerations of the brake pad holders together with the brake pads were recorded. Figure 8 presents the relation of the instantaneous vibration accelerations as a function of the braking time for the brake pad holders together with the brake pads on the side of the piston rod and the brake cylinder.

Then, from the point measures, the RMS effective value of the vibration accelerations was determined according to [38] described with relation (2):

$$A_{RMS} = \sqrt{\frac{1}{T}\int_0^T \left[a(t)\right]^2 dt} \tag{2}$$

where: T – averaging time in [s],
$a(t)$ – instantaneous value of the vibration accelerations in [m/s^2].

Then, applying the relation (2) from the instantaneous vibration accelerations for both pad holders, effective values were obtained for 40 instances of braking in the run-in process (speed 120 km/h). Figure 9 presents the ARMS relation for individual braking instances. From the experience gained during the test stand investigations of the disc brake, the authors know that the run-in of the pads is carried out for 25-30 braking instances after which (in conformity with [42]) over 75% of the surface is properly run in.

When analyzing the graph presented in Figure 9, one may observe that during subsequent braking instances performed under the same initial conditions (velocity, pressure, mass to be decelerated and disc

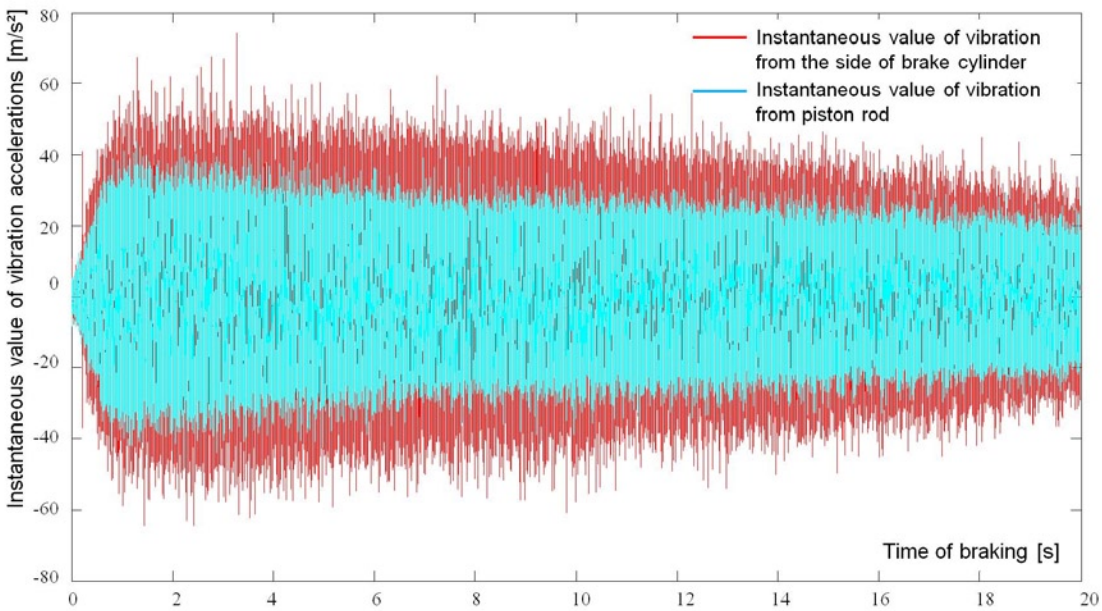


Fig. 8. Instantaneous value of the vibration accelerations of the brake pad holders during the first 20 seconds of braking from the speed of v=120 km/h, at the pressure of the pads on the disc N=25 kN and mass to be decelerated of M=5.7 tons

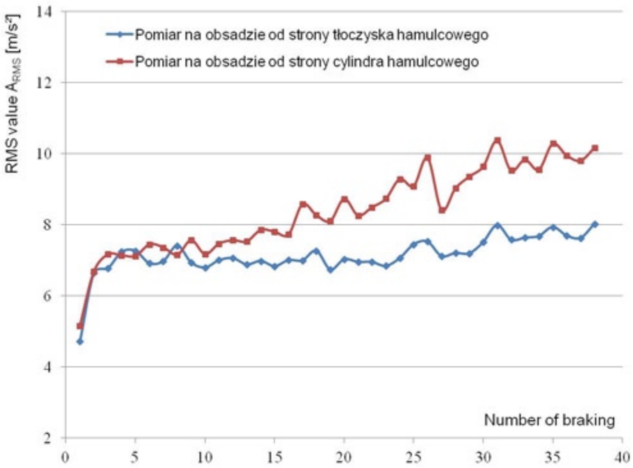


Fig. 9. Variability of the effective value ARMS of vibration accelerations for 40 braking instances during brake pad running-in

Next, an analysis of the signals of vibration accelerations in the frequency domain was performed. Figure 10 presents the spectrums of the vibration of the brake pad holders together with the brake pads (on both sides of the brake disc).

In the first place, for various braking (stopping and constant power), the frequency bands associated with the change in wear of the friction pads in the range of their thickness from 15 to 35 mm are determined, which directly influences the change in the geometry of the lever system. The α_1 (Fig. 1c) and α_2 (Fig. 3b and 3c) Angles change at the same time as the wear pads are worn in the lever mechanism. Frequency analysis has shown that in the 4600-4800 Hz frequency band is observed to change the effective vibration acceleration of friction pads both on the right side of the brake disc (brake cylinder side) and to the left of the disc, i.e. From the piston rod. For this frequency range, the dynamics of the change described by the dependence (3) in the vibrations of the pads new to the worn-out exceeds 6 dB. Table 3 shows the ARMS values for each of the vibration frequency bands of the friction pads of different thicknesses on the two sides of the brake disc during braking. Figure 11 graphically illustrates the dependence

temperature), the difference in the vibration of the left and right pad holder increase reaching the highest value after the pads and the disc run in. However, on the side of the brake cylinder, the increment of vibration in subsequent braking instances is lower compared to the other side of the lever mechanism (the side of the piston rod). Besides, it was observed that, in the process of running in of the pads, the effective value of the accelerations increases, which is to be explained by the fact that in the first instances of braking not entire surface of the brake pads is pressed against the disc. When the pads are run in, after approx. 25 braking instances, the effective value of the vibration acceleration stabilizes on the level of approx. 10 m/s^2 for the pad holder on the side of the brake cylinder and approx. 8 m/s^2 for the holders together with the pads on the side of the piston rod.

When analyzing the graph presented in figure 9, it was observed that, according to relation 3 [28] as the brake pads run in with the brake discs, the dynamics of the changes of vibration between both sides of the brake discs increases. The dynamics of the changes increase from 1dB in the first braking instance to 3 dB after 30 braking instances at fully run in brake pads of the disc brake:

$$D = 20\lg\left(\frac{A_2}{A_1}\right) \quad (3)$$

where: A1 – value of the point measure (ARMS) determined during braking on the brake pad holders on the side of the piston rod in [m/s^2],

A2 – value of the point measure (ARMS) determined during braking on the brake pad holders on the side of the cylinder in [m/s^2].

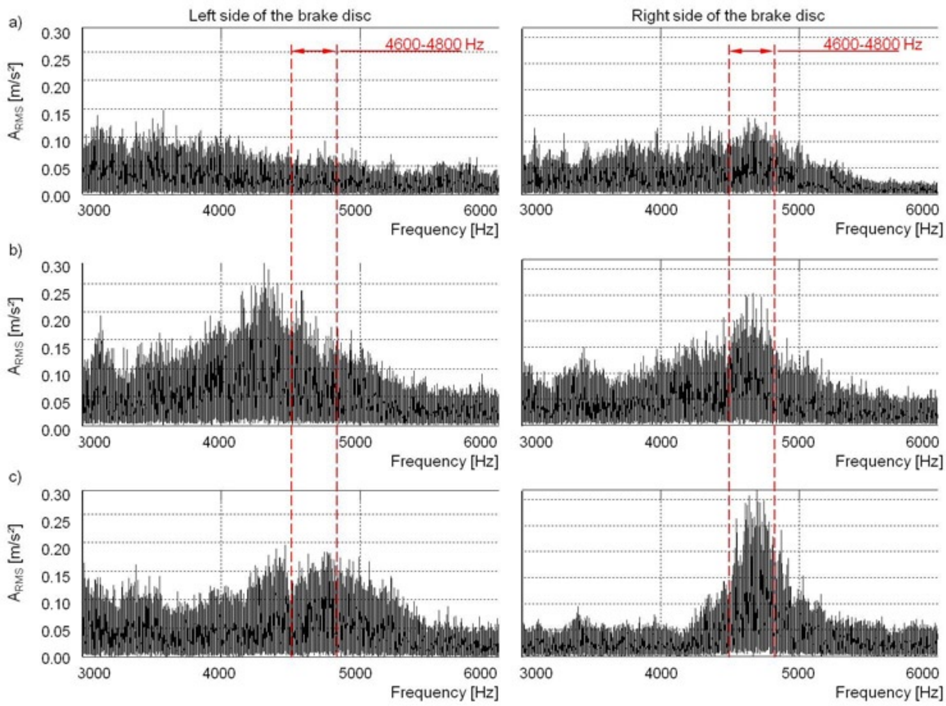


Fig. 10. ARMS dependence on frequency for two brake mounts with friction pads for the right and left sides of brake disc, a) for new pads, thicknesses G1=35 mm, b) for pads worn to thickness G2=25 mm, b) for pads worn to thickness G3=15 mm

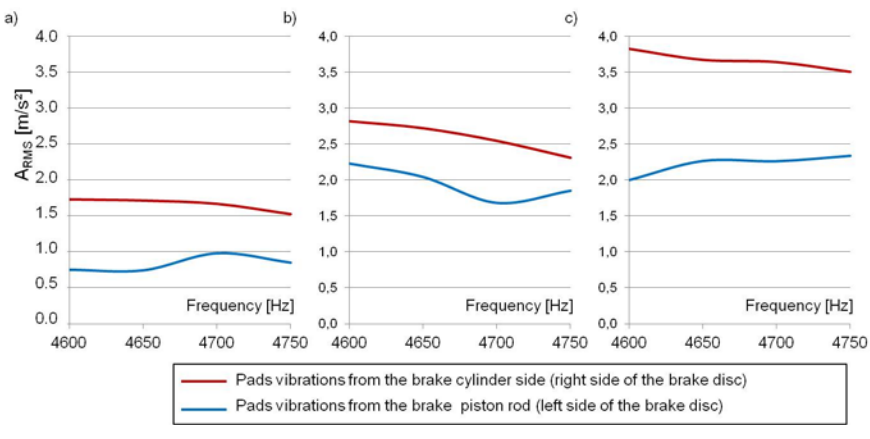


Fig. 11. ARMS relationship in the frequency band 4600-4750 Hz for two brake mounts with friction pads for the right and left of the brake disc, a) for new pads with a thickness of G1 = 35 mm, b) for pads worn up to a thickness of G2 = 25 mm, b) for pads worn up to a thickness of G3 = 15 mm

Table 3. Statement of the ARMS values of the pads vibrations on the right and left side of the brake disc for 4600-4800 Hz band for three cases of friction material thickness during braking

For new pads thicknesses G_1=35 [mm]				
Frequency [Hz]	ARMS-2 pads on the left side of the disc [m/s^2]	ARMS-1 pads on the right side of the disc [m/s^2]	Vibration difference [m/s^2]	Dynamics of change [dB]
4600-4650	0.754	1.728	0.974	7.208
4650-4700	0.745	1.712	0.967	7.226
4700-4750	0.978	1.667	0.688	4.630
4750-4800	0.850	1.524	0.673	5.066
For pads worn to thickness G_2=25 [mm]				
4600-4650	2.228	2.820	0.592	2.048
4650-4700	2.041	2.724	0.682	2.507
4700-4750	1.678	2.547	0.868	3.624
4750-4800	1.848	2.311	0.463	1.943
For pads worn to thickness G_3=35 [mm]				
4600-4650	1.997	3.825	1.828	5.647
4650-4700	2.263	3.672	1.408	4.203
4700-4750	2.259	3.641	1.381	4.145
4750-4800	2.335	3.506	1.171	3.531

of the values of vibration accelerations on the right and left sides of the brake disc during braking for three cases of friction material thickness in the frequency function for the 4600-4700 Hz band.

Analyzing the results of vibration of the friction pads on both sides of the brake disc during braking, it is found that the analysis of lining vibration acceleration in the frequency domain relative to the analysis in the field of amplitude allows observation of a greater difference in the vibration of the right and left pad of the brake disc. The dynamics of changes is in the range of 4.6-7.2 dB for new pads, 2-3.6 dB for worn pads up to 25 mm thick and 3.5-5.6 dB for pads worn up to 15 mm thick. In the case of amplitude analysis, the differences in right and left vibration of the disc are in the range of 1-3 dB depending on the surface condition of the friction pads.

The paper [30] presents the results of vibroacoustic tests of a railway disc brake in which it was shown that the 4600-4700 Hz frequency band enables the assessment of the disc condition of a railway brake. In contrast, pulse tests of the main disc brake assemblies showed that the 4400-4600 Hz band is a common resonance band of such elements as the disc, friction pads and a complex lever mechanism.

Based on the vibroacoustic tests in the field of amplitudes and time, it was found that the non-uniformity of the vibrations of the right and left friction pad is the result of disturbances in the braking process. It was shown that this is caused by the uneven mass distribution of the elements of the lever system of the right and left side of the brake disc and friction in the pin joints of the mechanism. Consequently, the disc is not pressed by the pads with equal force, which would result from the design and assumptions of a symmetrical brake lever mechanism. In a broader sense, the demonstrated irregularity in the pressure of the brake pads on the disc may also affect the extension of the braking distance.

8. Conclusion

The article presents the results of the author's tests on a certified brake station for testing brakes of rail vehicles in the field of tribology, thermovision and vibroacoustics. Based on these tests, it was proved that the brakes of the right and left of the disc brake friction pair were uneven. It was found that the propagation of unfavorable vibrations and noise through the lever system may also result from the geometry of the braking system. The conducted research proves that the current models of vibration and noise in braking systems are insufficient and take into account only the case of interaction of one friction element with the brake disc. Due to the design of the lever system, which should uniformly transfer the force from the brake cylinder to the friction pads, the friction phenomena and vibrations are different. It is caused by the uneven mass distribution of the elements of the right and left sides of the clamping mechanism, changes in the geometry of the system and friction in the kinematic nodes.

Uneven weight distribution of the braking system components relative to the right and left sides of the brake disc and friction in the bolt connections cause the following effects:

- - an increase in vibration acceleration by an average of 25% on the cast from the brake cylinder side relative to the cast from the brake piston rod side in the entire braking process based on analysis in the field of amplitude of acceleration of friction lining vibrations,
- - increase in vibration acceleration by approx. 50% on the cast from the brake cylinder side relative to the cast on the piston rod side in the 4600-4800 Hz frequency band by analyzing the vibration acceleration signals in the frequency domain,
- - increased wear of the friction linings on the brake cylinder side by approx. 10-18% compared to the friction linings on the brake piston rod side,
- - an increase in the average temperature of the friction linings by approx. 15-17% located on the side of the lighter brake cylinder relative to the heavier piston rod.

The paper shows that the unevenness in the mass distribution of the right and left side of the lever system is unfavorable from the point of the braking process. This results, differences in the value of forces acting on the disc. Vibrations on one side of the disc are intensified, which disturbs the braking process and may cause an increase in the braking distance.

Further work is planned to develop a model of a new lever mechanism, characterized by equal mass distribution of the right and left sides of the disc and a constant value of forces acting on the disc regardless of friction pad wear. Then, kinematic and dynamic analyses of the classic lever system (presented in the paper) in relation to the new one will be carried out, and an attempt will be made to carry out comparative tests on a brake stand.

The project was financed by National Centre for Research and Development, program LIDER V, contract No. LIDER/022/359/L-5/13/NCBR/2014

References

1. Babishin V, Taghipour S. An algorithm for estimating the effect of maintenance on aggregated covariates with application to railway switch point machines. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(4): 619-630, https://doi.org/10.17531/ein.2019.4.11.
2. Bowden FP, Leben L. The nature of sliding and the analysis of friction. Royal Society of London Proceedings Series A Mathematics Physics and Engineering Science 1939; 169: 371-391, https://doi.org/10.1098/rspa.1939.0004.
3. Brooks PC, Crolla DA, Lang AM, Schafer DR. Eigenvalue sensitivity analysis applied to disc brake squeal. In: Braking of Road Vehicles. Bury St. Edmunds, England: Institution of Mechanical Engineers, Mechanical Engineering Publications Limited, 1993: 135-143.
4. Earles SWE, Chambers PW. Disc brake squeal noise generation: predicting its dependency on system parameters including damping. International Journal of Vehicle Design 1987; 8: 538-552.
5. European Standard EN 14535-3:2015. Railway applications - Brake discs for railway rolling stock - Part 3: Brake discs, performance of the disc and the friction couple, classification: 12-16.
6. European Standard EN 14535-1:2005. Railway applications - Brake discs for railway rolling stock - Part 1: Brake discs pressed or shrunk onto the axle or drive shaft, dimensions and quality requirements: 22-23.
7. Fosberry RAC, Holubecki A. Disc brake squeal: its mechanism and suppression. Technical Report. Motor Industry Research Association 1961; 1: 34-36.
8. Ghazaly NM, El-Sharkawy M, Ahmed I. A Review of Automotive Brake Squeal Mechanisms. Journal of Mechanical Design and Vibration 2013; 1(1): 5-9
9. Gill A, Kadziński A. The determination procedure of the onset of the object wear-out period based on monitoring of the empirical failure intensity function. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(2): 282-287, https://doi.org/10.17531/ein.2015.2.16.
10. Glowacz A, Glowacz W, Kozik J, Piech K, Gutten M, Caesarendra W, Liu H, Brumercik F, Irfan M, Khan F. Detection of Deterioration of Three-phase Induction Motor using Vibration Signals. Measurement Science Review 2019; 19(6): 241-249, https://doi.org/10.2478/msr-2019-0031.
11. Glowacz A. Fault diagnosis of electric impact drills using thermal imaging. Measurement 2021; 171: 1-11, https://doi.org/10.1016/j.measurement.2020.108815.
12. Glowacz A. Acoustic fault analysis of three commutator motors. Mechanical Systems and Signal Processing 2019; 133: 1-17, https://doi.org/10.1016/j.ymssp.2019.07.007.
13. Hu Y, Wang S, Ai X. Research of the Vibration Source Tracking in Phase-Sensitive Optical Time-Domain Reflectometry Signals Based by Image Processing Method. Algorithms 2018; 11(8): 1-16, https://doi.org/10.3390/a11080117.
14. Jarvis RP, Mills B. Vibrations induced by friction. Proceedings of the Institution of Mechanical Engineers 1963; 178(32): 847-857, https://doi.org/10.1177/0020348363178001124.
15. Kinkaid NM, O'Reilly OM, Papadopoulos P. Automotive disc brake squeal. Journal of Sound and Vibration 2003; 267: 105-166, https://doi.org/10.1016/S0022-460X(02)01573-0.
16. Konowrocki R, Chojnacki A. Analysis of rail vehicles' operational reliability in the aspect of safety against derailment based on various methods of determining the assessment criterion. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(1): 73-85, https://doi.org/10.17531/ein.2020.1.9.
17. Lang AM, Smales H. An approach to the solution of disc brake vibration problems. in: Braking of Road Vehicles. Automobile Division of the Institution of Mechanical Engineers. Suffolk: Mechanical Engineering Publications Limited 1993: 223-231.
18. Meierhofer A, Hardwick C, Lewis R, Six K, Dietmaier P. Third body layer-expermental results and a model describing its influence on the traction coefficient. Wear 2014; 314: 148-154, https://doi.org/10.1016/j.wear.2013.11.040.
19. Merkisz J, Rymaniak Ł. The assessment of vehicle exhaust emissions referred to CO2 based on the investigations of city buses under actual conditions of operation. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(4): 522-529, https://doi.org/10.17531/ein.2017.4.5.
20. Millner N. An analysis of disc brake squeal. Technical Report 780332, SAE 1978: 22-24, https://doi.org/10.4271/780332.
21. Mills HR. Brake squeak. Technical Report 9000 B, Institution of Automobile Engineers 1938: 13-17.
22. Murakami H, Tsunada N, Kitamura T. A study concerned with a mechanism of disc-brake squeal. SAE Transactions 1984: 93(5): 604-616, https://doi.org/10.4271/841233.
23. North MR. Disc brake squeal, in: Braking of Road Vehicles. London: Automobile Division of the Institution of Mechanical Engineers, Mechanical Engineering Publications Limited 1976: 169-176.
24. Qu Y, Xie F, Su H, Meng G. Numerical analysis of stick-slip induced nonlinear vibration and acoustic responses of composite laminated plates with friction boundaries. Composite Structures 2021; 258: 1-12, https://doi.org/10.1016/j.compstruct.2020.113316.
25. Rudolph M, Popp K. The Influence of Contact Properties on Friction-Induced Brake Vibrations. Proceedings of the 3rd Contact Mechanics International Symposium 2001: 125-132, https://doi.org/10.1007/978-94-017-1154-8_14.
26. Rudolph M, Popp K. Brake squeal. In: Popp K. (Ed.), Detection, Utilization and Avoidance of Nonlinear Dynamical Effects in Engineering Applications. Final Report of a Join Research Sponsored by the German Federal Ministry of Education and Research, Shaker, Aachen 2001: 197-225.
27. Saumweber E. Auslegung und Leistungsgrenzen von Scheibenbremsen. ZEV-Glasses Annalen 1988; 112(4): 139-143.
28. Sawczuk W. Application of selected frequency characteristics of vibration signal for the evaluation of the braking process for railway disc brake. Diagnostyka 2015; 16(3): 33-38.
29. Sawczuk W. Application of vibroacoustic diagnostics to evaluation of wear of friction pads rail brake disc. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(4): 565-571, https://doi.org/10.17531/ein.2016.4.11.
30. Sawczuk W, Szymański GM. Diagnostics of the railway friction disc brake based on the analysis of the vibration signals in terms of resonant frequency. Archive of Applied Mechanics 2016; 86: 1-15.

31. Selech J, Andrzejczak K. An aggregate criterion for selecting a distribution for times to failure of components of rail vehicles. Eksploatacja i Niezawodnosc - Maintenance and Reliability, 2020; 22(1): 102-111, https://doi.org/10.17531/ein.2020.1.12.
32. Shen X, Lu L, Zeng D. Fatigue failure analysis of high strength bolts used for high-speed railway vehicle braking discs. Engineering Failure Analysis 2020; 115: 1-16, https://doi.org/10.1016/j.engfailanal.2020.104661.
33. Shi LB, Wang F, Ma L, Liu QY, Guo J, Wang WJ. Study of the friction and vibration characteristics of the braking disc/pad interface under dry and wet conditions. Tribology International 2018; 127: 553-544, https://doi.org/10.1016/j.triboint.2018.07.012.
34. Sinclair D. Frictional vibrations. Transactions of the American Society of Mechanical Engineers. Journal of Applied Mechanics 1955; 77: 207-213.
35. Singh S, Kumar R. Evaluation of human error probability of disc brake unit assembly and wheel set maintenance of Railway Bogie. Procedia Manufacturing 2015; 3: 3041-3048, https://doi.org/10.1016/j.promfg.2015.07.849.
36. Sinou JJ, Thouverez F, Jezequel L. Analysis of friction and instability by the centre manifoldtheory for a non-linear sprag-slip model. Journal of Sound and Vibration 2003; 265: 527-559,https://doi.org/10.1016/S0022-460X(02)01453-0.
37. Spurr RT. A theory of brake squeal. Proceedings of the Automobile Division. Institution of Mechanical Engineers 1961-1962; (1): 33-52, https://doi.org/10.1243/PIME_AUTO_1961_000_009_02.
38. Szymański GM, Tomaszewski F. Diagnostic of automatic compensators of valve clearance in combustion engine with the use of vibration signal. Mechanical Systems and Signal Processing 2016; 68-69: 479-490, https://doi.org/10.1016/j.ymssp.2015.07.015.
39. Świderski A, Borucka A, Jacyna-Gołda I, Szczepański E. Wear of brake system components in various operating conditions of vehicle in the transport company. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 1-9, https://doi.org/10.17531/ein.2019.1.1.
40. Tan Y, Zhang H, Zhou Y. A Simple-to-Implement Fault Diagnosis Method for Open Switch Fault in Wind System PMSG Drives without Threshold Setting. Energies 2018; 11(10): 1-18, https://doi.org/10.3390/en11102571.
41. Triches M, Gerges SNY, Jordan R. Analysys of brake squale noise using finite element method. Applied Acoustic 2008; 69(2): 147-162, https://doi.org/10.1016/j.apacoust.2007.10.003.
42. UIC Code 541-3: Brakes - Disc brakes and their application - General conditions for the approval of brake pads. 7th edition, January 2010.
43. Wang XC, Huang B, Wang RL, Mo JL, Ouyang H. Friction-induced stick-slip vibration and its experimental validation. Mechanical Systems and Signal Processing 2020; 142: 2-21, https://doi.org/10.1016/j.ymssp.2020.106705.
44. Zhao X, Gräbner N, Wagner U. Avoiding creep groan: Investigation on active suppression of stick-slip limit cycle vibrations in an automotive disk brake via piezoceramic actuators. Journal of Sound and Vibration 2019; 441: 174-186, https://doi.org/10.1016/j.jsv.2018.10.049.

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl
Article citation info:
Zhou S, Zhang J, You L, Zhang Q. Uncertainty propagation in structural reliability with implicit limit state functions under aleatory and epistemic uncertainties. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 231–241, http://doi.org/10.17531/ein.2021.2.3.

Uncertainty propagation in structural reliability with implicit limit state functions under aleatory and epistemic uncertainties

Indexed by: Web of Science Group

Shuang Zhou[a], Jianguo Zhang[b,*], Lingfei You[c], Qingyuan Zhang[d]

[a] School of Reliability and Systems Engineering, Beihang University, Xueyuan Road No.37, Haidian, District, Beijing 100191, China
[b] Science and Technology on Reliability and Environmental Engineering Laboratory, Beihang University, Xueyuan Road No.37, Haidian District, Beijing 100191, China
[c] Institute of Unmanned System, Beihang University, Xueyuan Road No.37, Haidian District, Beijing 100191, China
[d] School of Aeronautic Science and Engineering, Beihang University, Xueyuan Road No.37, Haidian District, Beijing 100191, China

Highlights

- A novel surrogate model is given for implicit function under uncertain random variable.
- The concepts of chance reliability and chance reliability index (CRI) are proposed.
- A new level-2 uncertainty propagation model is provided for uncertain random structure.
- The principles for choosing reasonable uncertainty propagation types are presented.

Abstract

Uncertainty propagation plays a pivotal role in structural reliability assessment. This paper introduces a novel uncertainty propagation method for structural reliability under different knowledge stages based on probability theory, uncertainty theory and chance theory. Firstly, a surrogate model combining the uniform design and least-squares method is presented to simulate the implicit limit state function with random and uncertain variables. Then, a novel quantification method based on chance theory is derived herein, to calculate the structural reliability under mixed aleatory and epistemic uncertainties. The concepts of chance reliability and chance reliability index (CRI) are defined to show the reliable degree of structure. Besides, the selection principles of uncertainty propagation types and the corresponding reliability estimation methods are given according to the different knowledge stages. The proposed methods are finally applied in a practical structural reliability problem, which illustrates the effectiveness and advantages of the techniques presented in this work.

Keywords

structural reliability, uncertainty quantification, uncertainty propagation, reliability index, uncertainty theory.

1. Introduction

Structural reliability assessment has been widely recognized as vital in engineering product design and development [7]. In the context of structural reliability assessment, uncertainty propagation plays a significant role, which aims to quantify the uncertainties of input factors and calculate the overall uncertainty within the model response in reliability estimation [36].

Before propagating the structure's uncertainty, a primary issue is to choose a reasonable mathematical theory related to the types of uncertainty, to quantify the uncertainty [12,38]. In practical structural engineering problems, uncertainty can be divided into two categories: aleatory uncertainty derived from inherent randomness of physical behavior, while the epistemic uncertainty arising out of lack of knowledge [10].

Probability theory is regarded as the most effective tools to describe aleatory uncertainty in structural reliability assessment. Over the last decades, numerous reliability assessment methods based on probability theory have been developed, including first-order reliability method (FORM) [23], second-order fourth moment [29] Monte Carlo simulation (MCS) [24], FORM-sampling simulation method [22], envelope function method [28], response surface method (RSM) [6], and Bayesian networks method [26]. Although these probabilistic methods typically make sense in uncertainty quantification and propagation when the structure is mainly affected by aleatory uncertainty, they do not work well in the scenarios involving great epistemic uncertainty [37]. For example, the distribution of input factors may not be precisely obtained due to insufficient sample data. Consequently, several alternative non-probabilistic theories have been developed to describe the epistemic uncertainty in reliability assessment.

The general non-probabilistic structural reliability assessment theories consist of fuzzy set theory [9], fuzzy random theory [13], possibility theory [1], interval theory [5, 27], and evidence theory [39]. The fuzzy and possibility measures fail to satisfy the duality property, which will make it difficult for decision-makers to understand the results [33]. Moreover, interval and evidence theories will lead to an over-conservative result due to the interval extension problems [38]. To overcome the shortcomings of the above-mentioned theories, a new mathematical framework called uncertainty theory was introduced to deal with epistemic uncertainty.

(*) Corresponding author.
E-mail addresses: S. Zhou - zsdyx88@163.com, J. Zhang - zjg@buaa.edu.cn, L. You - by1514119@buaa.edu.cn, Q. Zhang - zhangqingyuan@buaa.edu.cn

Uncertainty theory proposed by Liu [18] in 2007 to describe the belief degree of human, has been successfully applied in various areas such as decision making [30], uncertain insurance [19, 32], uncertain risk and reliability analysis [34, 35]. Uncertainty theory is considered a reasonable and useful tool to express epistemic uncertainty, compared with the theories mentioned above [12]. Since the uncertain measure satisfies the axiom of duality, normality, and subadditivity, the results produced by the uncertainty theory are more in line with real engineering conditions [8]. Hence, in this work, uncertainty theory is chosen to express epistemic uncertainty and describe human thinking processes. In practical structural problems, there are usually two types of input factors that embody different types of uncertainties at the same time. Some input factors may suffer great epistemic uncertainty and are described by uncertainty theory, while some others may be primarily determined by aleatory uncertainty and are modeled based on probability theory. These structures comprising both aleatory and epistemic uncertainties are called uncertain random structures in this paper. It is impossible to analyze the reliability of uncertain random structures only by probability theory or uncertainty theory [38].

To solve this problem, chance theory was established by Liu [20] in 2013 to propagate aleatory and epistemic uncertainties together. Chance theory can be understood as a combination of probability theory and uncertainty theory, which also satisfies normality, duality, and subadditivity theorems. In recent years, chance theory has been successfully used in various fields such as project scheduling [11], uncertain random risk analysis [8], uncertain random programming [25], and systems reliability analysis [31, 38]. Especially, a hybrid model of structural reliability analysis based on chance theory was proposed by Zhang [37] in 2019, and a new reliability index was proposed. However, this method has the following disadvantages. Firstly, there is no corresponding reliability analysis method when the implicit limit state function (LSF) contains both random and uncertain variables. Secondly, the defined reliability and reliability index do not involve time dynamic parameters. Thirdly, this reliability analysis method does not consider the problem of uncertainty propagation.

For completeness, this paper uses a uniform design (UD) combined with the least-squares (LS) method to simulate LSF, which adapts to both random and uncertain variables. The UD is a novel kind of experimental design method founded by Fang and Wang [3], defined according to the uniform distribution in number theory [40]. Compared to the orthogonal design (OD), factorial design (FD), and Latin hypercube sampling (LHS) methods, the UD method appears to be more advanced if the number of experimental factors is large and the number of experiments is limited [4].

Besides the above research, this paper's main contribution is to provide a new uncertainty propagation method for structural reliability assessment. Uncertainty propagation aims to estimate structural output responses by propagating the input factors essential for structural reliability assessment and safety design [36]. Normally, uncertainty propagation can be classified into the form of level-1 and level-2 [14]. For level-1 propagation, the values of input factors can be characterized by epistemic or aleatory uncertainties at the same level [2]. For level-2 propagation, the values of input factors are represented by aleatory uncertainties on the first level. Epistemic uncertainties describe the parameters of probability distributions in the second level [34]. These two types of uncertainty propagation methods are commonly used in risk assessment. Comprehensive research about this was reported by Hu et al. [8], who presented a framework for propagation methods corresponding to different knowledge stages in fault tree analysis. However, there are no literature about the level-2 uncertainty propagation modeling and propagation type selection methods for structural reliability assessment. Hence, this paper aims to develop some propagation analysis methods and the principles for the selection of propagation type in structural reliability assessment.

The remainder of this work is organized as follows: Section 2 briefly discusses some important mathematical concepts of uncertainty and chance theory. A new surrogate model combining UD and LS method is proposed for implicit LSF in Section 3. Section 4 provides a novel structural reliability quantification model based on chance theory. Some principles for choosing appropriate uncertainty propagation types are discussed, and corresponding reliability calculation methods are provided in Section 5. In Section 6, a practical engineering case study is carried out to show the proposed method's rationality. Finally, some conclusions are presented in Section 7.

2. Preliminaries

In this section, some fundamental knowledge and results regarding the uncertainty theory and chance theory are introduced.

2.1. Uncertainty theory

Uncertainty theory is a fairly new branch of axiomatic mathematics, and has been widely applied in various areas. In the uncertainty theory, the human belief degree of events are quantified by defining uncertain measures.

Definition 2.1 (Uncertain measure [15]) Let Γ be a nonempty set, and $\mathcal{L}$ be a σ- algebra over Γ. Each element Λ in $\mathcal{L}$ is called an event. Then, a set function M is defined as an uncertain measure if it satisfies normality, duality, and subadditivity axioms.

Definition 2.2 (Uncertain variable [18]) An uncertain variable is a measurable function τ from an uncertainty space $(\Gamma, \mathcal{L}, \text{M})$ to the set of real numbers, i.e., $\{\tau\in \text{B}\}$ is an event for any Borel set B of real numbers.

Definition 2.3 (Uncertainty distribution [15]) The uncertainty distribution $\Phi(x)$ of an uncertain variable τ can be defined by $\Phi(x) = \text{M}\{\tau\le x\}$ for any real number x.

A regular uncertainty distribution $\Phi(x)$ is defined as an uncertainty function that is continuous and strictly increasing with respect to x.

Example 2.1 An uncertain variable τ is defined as a normal uncertain variable if it has a normal uncertainty distribution:

$$\Phi(x) = (1+\exp(\frac{\pi(m-x)}{\sqrt{3}\sigma}))^{-1} \quad x\in \text{R} \tag{1}$$

It is denoted by $\tau \sim \mathcal{N}(m, \sigma)$, where m is the expected value and σ is the standard deviation.

Example 2.2 An uncertain variable τ is defined as linear variable if it has a linear uncertainty distribution:

$$\Phi(x) = \begin{cases} 0, & \text{if } x \le a \\ \dfrac{x-a}{b-a}, & \text{if } a < x \le b \\ 1, & \text{if } b < x \end{cases} \tag{2}$$

It is denoted by $\tau \sim \mathcal{L}(a, b)$, where a and b are real numbers with $a < b$.

Since the uncertainty theory can describe the incomplete information contained in design variables, the epistemic uncertainty (especially human) can be characterized by uncertain variables and uncertainty distribution in the uncertainty space [16, 17].

Definition 2.4 (Inverse uncertainty distribution [15]) Let τ be an uncertain variable with regular uncertain distribution $\Phi(x)$. The inverse function $\Phi^{-1}(u)$ is known as the inverse uncertainty distribution of τ.

Theorem 2.1 (Operational law [18]) Let $\tau_1, \tau_2, \cdots, \tau_n$ be independent uncertain variables with regular uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If $f(\tau_1, \tau_2, \cdots, \tau_n)$ is continuous, strictly increasing with respect to $\tau_1, \tau_2, \cdots, \tau_m$ and strictly decreasing with respect to $\tau_{m+1}, \tau_{m+2}, \cdots, \tau_n$, then $\tau = f(\tau_1, \tau_2, \cdots, \tau_n)$ is an uncertain variable with inverse uncertainty distribution:

$$\Psi^{-1}(u)=f(\Phi_1^{-1}(u),\Phi_2^{-1}(u),\cdots,\Phi_m^{-1}(u),\Phi_{m+1}^{-1}(1-u),\Phi_{m+2}^{-1}(1-u),\cdots,\Phi_n^{-1}(1-u)) \tag{3}$$

2.2. Chance theory

As a combination of uncertainty and probability theory, the chance theory is applied as a new tool to deal with problems affected by both uncertainty and randomness. The basic concept involves the chance measure of an uncertain random event in a chance space.

Definition 2.5 (Chance measure [20]) Let $(\Gamma, \mathcal{L}, \mathrm{M}) \times (\Omega, \mathcal{A}, \Pr)$ be a chance space, and let $\Theta \in \mathcal{L} \times \mathcal{A}$ be an event. Then the chance measure of Θ can be defined as:

$$\mathrm{Ch}\{\Theta\}=\int_0^1 \Pr\{\omega\in\Omega \mid \mathrm{M}\{\gamma\in\Gamma \mid (\gamma,\omega)\in\Theta\}\ge r\}\mathrm{d}r \tag{4}$$

Theorem 2.2 Let $(\Gamma, \mathcal{L}, \mathrm{M}) \times (\Omega, \mathcal{A}, \Pr)$ be a chance space, then the chance measure $\mathrm{Ch}\{\Lambda\times\mathrm{A}\}=\mathrm{M}\{\Lambda\}\times\Pr\{\mathrm{A}\}$ for any $\Lambda\in\mathcal{L}$ and any $\mathrm{A}\in\mathcal{A}$. Especially, we have $\mathrm{Ch}\{\varnothing\}=0$, $\mathrm{Ch}\{\Gamma\times\Omega\}=1$ [15].

Definition 2.6 (Uncertain random variable [20]) An uncertain random variable is a measurable function ξ from a chance space $(\Gamma, \mathcal{L}, \mathrm{M}) \times (\Omega, \mathcal{A}, \Pr)$ to the set of real numbers such that $\{\xi\in\mathrm{B}\}$ is an event in $\mathcal{L} \times \mathcal{A}$ for any Borel set B of real numbers.

Definition 2.7 (Chance distribution [15]) Let ξ be an uncertain random variable. Then its chance distribution is defined as $\Phi(x)=\mathrm{Ch}\{\xi\le x\}$ for any real number x.

Theorem 2.3 Let $\eta_1,\eta_2,\cdots,\eta_m$ be independent random variables with probability distributions $\Psi_1,\Psi_2,\cdots,\Psi_m$, and let $\tau_1,\tau_2,\cdots,\tau_n$ be independent uncertain variables with uncertainty distributions $\Upsilon_1,\Upsilon_2,\cdots,\Upsilon_n$, respectively. If f is a measurable function, then the uncertain random variable $\xi=f(\eta_1,\eta_2,\cdots,\eta_m,\tau_1,\tau_2,\cdots,\tau_n)$ has a chance distribution [20]:

$$\Phi(x)=\int_{\Re^m} F(x;y_1,y_2,\cdots,y_m)\mathrm{d}\Psi_1(y_1)\mathrm{d}\Psi_2(y_2)\cdots\mathrm{d}\Psi_m(y_m) \tag{5}$$

where $F(x;y_1,y_2,\cdots,y_m)$ is the uncertainty distribution of the uncertain variable $f(y_1,y_2,\cdots,y_m,\tau_1,\tau_2,\cdots,\tau_n)$ for the any real numbers $y_1,y_2,\cdots,y_m$.

Besides, assume f is continuous, strictly increasing with respect to $\tau_1,\tau_2,\cdots,\tau_k$ and strictly decreasing with respect to $\tau_{k+1},\tau_{k+2},\cdots,\tau_n$. Then $F(x;y_1,y_2,\cdots,y_m)$ is the root u of the following equation:

$$f(y_1,y_2,\cdots,y_m,\Upsilon_1^{-1}(u),\cdots,\Upsilon_k^{-1}(u),\Upsilon_{k+1}^{-1}(1-u),\cdots,\Upsilon_n^{-1}(1-u))=x \tag{6}$$

Theorem 2.4 (Expected value [21]) Let $\eta_1,\eta_2,\cdots,\eta_m$ be independent random variables with probability distributions $\Psi_1,\Psi_2,\cdots,\Psi_m$, and let $\tau_1,\tau_2,\cdots,\tau_n$ be independent uncertain variables with regular uncertainty distributions $\Upsilon_1,\Upsilon_2,\cdots,\Upsilon_n$, respectively. If f is continuous and strictly increasing with respect to $\tau_1,\tau_2,\cdots,\tau_k$ and strictly decreasing with respect to $\tau_{k+1},\tau_{k+2},\cdots,\tau_n$. Then the uncertain random variable $\xi=f(\eta_1,\eta_2,\cdots,\eta_m,\tau_1,\tau_2,\cdots,\tau_n)$ has an expected value:

$$\begin{aligned}E[\xi]=\int_{\Re^m}\int_0^1 f(&y_1,y_2,\cdots,y_m,\Upsilon_1^{-1}(u),\cdots,\Upsilon_k^{-1}(u),\\ &\Upsilon_{k+1}^{-1}(1-u),\cdots,\Upsilon_n^{-1}(1-u))\mathrm{d}u\mathrm{d}\Psi_1(y_1)\cdots\mathrm{d}\Psi_m(y_m)\end{aligned} \tag{7}$$

Theorem 2.5 (Variance [15]) Let $\eta_1,\eta_2,\cdots,\eta_m$ be independent random variables with probability distributions $\Psi_1,\Psi_2,\cdots,\Psi_m$, and let $\tau_1,\tau_2,\cdots,\tau_n$ be independent uncertain variables with regular uncertainty distributions $\Upsilon_1,\Upsilon_2,\cdots,\Upsilon_n$, respectively. Assuming f is continuous, strictly increasing with respect to $\tau_1,\tau_2,\cdots,\tau_k$ and strictly decreasing with respect to $\tau_{k+1},\cdots,\tau_n$, then $\xi=f(\eta_1,\eta_2,\cdots,\eta_m,\tau_1,\tau_2,\cdots,\tau_n)$ has a variance:

$$\begin{aligned}V[\xi]=2\int_{\Re^m}\int_0^{+\infty} &x(1-F(e+x;y_1,\cdots,y_m)+F(e-x;y_1,\cdots,y_m))\\ &\mathrm{d}x\mathrm{d}\Psi_1(y_1)\cdots\mathrm{d}\Psi_m(y_m)\end{aligned} \tag{8}$$

where e is the expected value $E[\xi]$ of ξ, and $F(x;y_1,\cdots,y_m)$ is the uncertainty distribution of uncertain variable $f(y_1,y_2,\cdots,y_m,\tau_1,\tau_2,\cdots,\tau_n)$ for any real numbers $y_1,y_2,\cdots,y_m$, which is also the root of Equation (6).

3. Advanced UD-LS surrogate model for implicit limit state functions

In the practical structural reliability problems, the analytical expression of LSF is generally unknown. The traditional RSM of structural reliability analysis is iteratively obtained based on the probabilistic reliability index (PRI). Furthermore, the traditional RSM is only suitable for random variables and requires a large number of test sample data. Thus, a new surrogate model is established by combining UD with the LS method considering both of aleatory and epistemic information.

The structure's response is obtained by experiment or finite element analysis, and the sample points used to fit the surrogate model are determined by the design of experiments (DOE) methods. Compared with traditional DOE methods, the UD method is more stable and efficient [4]. UD can maintain the results with high stability and accuracy even with a small sample data. Similar to the OD approach, the UD method can be used to generate experiment points by a series of designed UD tables. The representation of a specific UD table is $U_n(m^n)$ or $U_n^*(m^n)$, where U denotes the UD table, m represents the number of levels and the number of experiments required, n is the number of input factors, and $*$ represents the UD table with a smaller deviation and better uniformity. This work presents only a brief introduction of the UD method, and interested readers can refer to relevant research literature [3, 4, 40]. The quadratic polynomial surrogate model without the cross-terms is chosen as the response surface function of the structure.

$$\hat{f}(\boldsymbol{x})=b_0+\sum_{i=1}^{n} b_i x_i+\sum_{j=n+1}^{2n} b_j x_j^2 \tag{9}$$

where $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ is the vector of input factors, x_i is a random variable or an uncertain variable. $\boldsymbol{b}=(b_0,b_1,\cdots,b_{2n})^{\mathrm{T}}$ is $2n+1$ undetermined coefficients vector in the surrogate model [4]. According to the LS approach, $\boldsymbol{b}$ can be estimated based on $\boldsymbol{b}=(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})^{-1}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{y}$, where $\boldsymbol{a}$ is the regression coefficients vector with $m\times(2n+1)$ orders, $\boldsymbol{y}=(f(x_1),f(x_2),\cdots,f(x_m))^{\mathrm{T}}$ is the real responses vector of the structure.

Some indexes are used for validation to verify the surrogate model's fitting performance and check the accuracy. Among them, the coefficient of determination R^2 is the most crucial measurement index:

$$R^2=1-\frac{\sum_{i=1}^{m}(\hat{f}_i(\boldsymbol{x})-f_i(\boldsymbol{x}))^2}{\sum_{i=1}^{m}(f_i(\boldsymbol{x})-\bar{f}(\boldsymbol{x}))^2}\quad (0\le R^2\le 1) \tag{10}$$

where $\bar{f}(\boldsymbol{x})$ is the expected value of all the real responses $f_i(\boldsymbol{x})$, and $\hat{f}_i(\boldsymbol{x})$ are the simulation values of the responses. The closer the value of R^2 to 1, the higher is the accuracy of UD-LS surrogate model fitting.

According to the stress-strength interference model and the UD-LS surrogate model, the LSF $G(\boldsymbol{x},\alpha)$ of a structural system under mixed aleatory and epistemic uncertainties can be expressed as:

$$G(\boldsymbol{x},\alpha) = S_{\text{threshold}} - \hat{f}(\boldsymbol{x},\alpha) \tag{11}$$

where $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ is the input factor that affects the structural functioning, $S_{\text{threshold}}$ is the allowable threshold of structural response, and α is a dynamic input parameter associated with time.

4. Structural reliability assessment method under mixed aleatory and epistemic uncertainties

In a complex structural system, some design variables may have enough samples for estimating their probability distribution, which can be described by random variables. Nonetheless, some other design variables may lack sufficient data, which can be estimated by domain experts and regarded as uncertain variables. The structure cannot be simply considered to be a random or uncertain structure model under mixed aleatory and epistemic uncertainties [37]. This section put forward an advanced structural reliability assessment method for this issue depending on chance measure and belief reliability theory.

4.1. Uncertainty quantification for structural reliability based on chance measure

Let $(\Gamma, \mathcal{L}, \mathrm{M}) \times (\Omega, \mathcal{A}, \Pr)$ be a chance space, and the LSF of structure contains uncertain random input factors $x_1, x_2, \cdots, x_n$. In the present work, the input factors are uniformly described by uncertain random variables $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$, then the chance reliability of structure based on the chance measure can be defined as follows.

Definition 4.1 Assuming that $G(\xi,\alpha)$ is the LSF of a structure, in which ξ is the vector of uncertain random variables, the chance reliability is defined as the chance measure of the reliability event $\{G(\xi,\alpha) > 0\}$ at α:

$$\mathrm{Ch}_{\text{reliability}}(\alpha) = \mathrm{Ch}\{G(\xi,\alpha) > 0\} \tag{12}$$

Because of the duality of chance measure, the chance measure of a failure event $\{G(\xi,\alpha) \le 0\}$ at α can be derived as:

$$\begin{aligned}\mathrm{Ch}_{\text{failure}}(\alpha) &= \mathrm{Ch}\{G(\xi,\alpha) \le 0\} \\ &= 1 - \mathrm{Ch}\{G(\xi,\alpha) > 0\} \\ &= 1 - \mathrm{Ch}_{\text{reliability}}(\alpha)\end{aligned} \tag{13}$$

Consequently, the uncertainty of a safety event at α in structure can be quantified by $\mathrm{Ch}_{\text{reliability}}(\alpha)$ with a numerical value of $[0, 1]$. $\mathrm{Ch}_{\text{failure}}(\alpha)$ describes the confidence how a failure even will be happened at α. Obviously, the higher the $\mathrm{Ch}_{\text{failure}}(\alpha)$, more is the possibility that the failure event will occur at α. The theorem to be defined below provides computational methods for practical engineering applications.

Theorem 4.1 Let the LSF of a structure contain independent random variables $\eta_1, \eta_2, \cdots, \eta_p$ with probability distributions $\Psi_1, \Psi_2, \cdots, \Psi_p$, and independent uncertain variables $\tau_1, \tau_2, \cdots, \tau_q$ with regular uncertainty distributions $\Upsilon_1, \Upsilon_2, \cdots, \Upsilon_q$, respectively. If the LSF $G(\eta_1, \cdots, \eta_p;\ \tau_1, \cdots, \tau_q;\ \alpha)$ is continuous and strictly increasing with respect to $\tau_1, \tau_2, \cdots, \tau_k$, and strictly decreasing with respect to $\tau_{k+1}, \tau_{k+2}, \cdots, \tau_q$, then the chance reliability of the structural system at α can be rewritten as:

$$\mathrm{Ch}_{\text{reliability}}(\alpha) = \int_{\Re^p} F(0; y_1, y_2, \cdots, y_p; \alpha)\mathrm{d}\Psi_1(y_1)\mathrm{d}\Psi_2(y_2)\cdots\mathrm{d}\Psi_p(y_p) \tag{14}$$

where $F(0; y_1, y_2, \cdots, y_p; \alpha)$ is the root u of the following equation for any real numbers $y_1, y_2, \cdots, y_p$:

$$G(y_1, y_2, \cdots, y_p;\ \Upsilon_1^{-1}(1-u), \cdots, \Upsilon_k^{-1}(1-u), \Upsilon_{k+1}^{-1}(u), \cdots, \Upsilon_q^{-1}(u);\ \alpha) = 0 \tag{15}$$

Proof. According to the Theorem 2.3 and Definition 4.1, the chance reliability can be computed as follows:

$$\begin{aligned}\mathrm{Ch}_{\text{reliability}}(\alpha) &= \mathrm{Ch}\{G(\eta_1, \cdots, \eta_p;\ \tau_1, \cdots, \tau_q;\ \alpha) > 0\} \\ &= \int_{\Re^p} \mathrm{M}\{G(y_1, \cdots, y_p;\ \tau_1, \cdots, \tau_q;\ \alpha) > 0\}\mathrm{d}\Psi_1(y_1)\cdots\mathrm{d}\Psi_p(y_p)\end{aligned} \tag{16}$$

where $\mathrm{M}\{G(y_1, y_2, \cdots, y_p;\ \tau_1, \cdots, \tau_q;\ \alpha) > 0\} = F(0; y_1, y_2, \cdots, y_p; \alpha)$ is the root u of Equation (15).

The proof is completed.

4.2. The new chance reliability index based on uncertain random variables

PRI in probability space is a vital indicator for quality of structure, and it can be used to describe the structural reliability under aleatory uncertainty. However, the PRI cannot accurately measure the reliability under mixed aleatory and epistemic uncertainties. For completeness, a novel chance reliability index (CRI) is defined using the expected value and variance of the uncertain random variable, showing the reliable degree of a structure in chance space.

Definition 4.2 Let $(\Gamma, \mathcal{L}, \mathrm{M}) \times (\Omega, \mathcal{A}, \Pr)$ be an chance space, the LSF $G(\xi,\alpha) = G(\eta_1, \cdots, \eta_p; \tau_1, \cdots, \tau_q; \alpha)$ of a structure contains independent random variables $\eta_1, \eta_2, \cdots, \eta_p$ with probability distributions $\Psi_1, \Psi_2, \cdots, \Psi_p$, and independent uncertain variables $\tau_1, \tau_2, \cdots, \tau_q$ with regular uncertainty distributions $\Upsilon_1, \Upsilon_2, \cdots, \Upsilon_q$, respectively. Then the CRI of the structural system at α can be given as follows:

$$\beta_{\text{chance}}(\alpha) = \frac{E[G(\xi,\alpha)]}{\sqrt{V[G(\xi,\alpha)]}} \tag{17}$$

where $E[G(\xi,\alpha)]$ is the expected value, $V[G(\xi,\alpha)]$ is the variance of LSF, and $\xi = \xi_1, \xi_2 \cdots \xi_{p+q}$ is an uncertain random vector.

If the LSF is continuous and strictly increasing with respect to $\tau_1, \tau_2, \cdots, \tau_k$, and strictly decreasing with respect to $\tau_{k+1}, \tau_{k+2}, \cdots, \tau_q$, then according to the theorems 2.4 and 2.5, the expected value and variance of LSF at α can be calculated as:

$$\begin{aligned}E[G(\xi,\alpha)] = \int_{\Re^p}\int_0^1 &G(y_1, \cdots, y_p; \Upsilon_1^{-1}(u), \cdots, \Upsilon_k^{-1}(u), \\ &\Upsilon_{k+1}^{-1}(1-u), \cdots, \Upsilon_q^{-1}(1-u); \alpha)\mathrm{d}u\mathrm{d}\Psi_1(y_1)\cdots\mathrm{d}\Psi_p(y_p)\end{aligned} \tag{18}$$

$$\begin{aligned}V[G(\xi,\alpha)] = 2\int_{\Re^p}\int_0^{+\infty} &x(1 - F(e_{G(\xi,\alpha)} + x; y_1, \cdots, y_p; \alpha) + F(e_{G(\xi,\alpha)} - x; y_1, \cdots, y_p; \alpha)) \\ &\mathrm{d}x\mathrm{d}\Psi_1(y_1)\cdots\mathrm{d}\Psi_p(y_p)\end{aligned} \tag{19}$$

where $e_{G(\xi,\alpha)} = E[G(\xi,\alpha)]$, and $F(x; y_1, \cdots, y_p; \alpha)$ is the root u of the following equation for any real numbers $y_1, y_2, \cdots, y_p$:

$$G(y_1, y_2, \cdots, y_p; \Upsilon_1^{-1}(u), \cdots, \Upsilon_k^{-1}(u), \Upsilon_{k+1}^{-1}(1-u), \cdots, \Upsilon_q^{-1}(1-u); \alpha) = x \tag{20}$$

The chance-measure-based CRI is offered as a tool to measure the confidence that a reliability event will occur in the structural system affected by both aleatory and epistemic uncertainties. A larger $\beta_{\text{chance}}(\alpha)$ indicates a better possibility that the reliability event will occur.

5 Joint uncertainty propagation method for structural reliability assessment

Uncertainty propagation plays a significant role in reliability problem, which aims to estimate structural output responses by propagating the input factors essential for reliability assessment and safety design. To make it possible for decision-makers to find an appropriate uncertainty propagation types under different knowledge stages, a new joint uncertainty propagation technique is presented in this section. Therefore, the selection principles of uncertainty propagation types are developed in section 5.1. Section 5.2 briefly introduces the level-1 uncertainty propagation method, especially the propagation of uncertain random structure. The novel level-2 joint uncertainty propagation method for structural reliability assessment is proposed in Section 5.3.

5.1. The principles of uncertainty propagation types selection

In general, uncertainty propagation can be classified into the form of level-1 and level-2. The quantification and propagation of uncertainty runs through the whole analysis process. To explain the uncertainty propagation of level-1 and level-2 types, the probability theory is utilized to express aleatory uncertainties, while the uncertainty theory is used to describe the epistemic uncertainties. $G(\boldsymbol{x})$ is assumed to be the LSF established in Section 3, where $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ is the input factors vector, and G is the output. To analyze the uncertainty of output G, the uncertainty expressions of the input factors needs to be studied, in addition to their propagations through LSF.

According to the knowledge stage of reliability analyst, the reliability evaluation types can divided into different stages. For example, uncertainty propagation types can be divided into five different stages, shown in Fig.1. At stage 1, the reliability analyst has no sample data on $x_1, x_2, \cdots, x_n$. So uncertain variables are used to describe all the input factors. In this situation, the uncertainty propagations are in level-1 type. At stage 2, the reliability analyst collects more sample data and improves his knowledge. The distribution function type of x_1 is known, which is the probability distribution type. Nonetheless, the shaping parameters of this probability distribution are still lacking and can be described by uncertain variables. $x_2, x_3 \cdots, x_n$ are still described by uncertain variables. In this case, the uncertainty propagations will be in level-2 type. At stage 3, the knowledge of reliability analyst improves further. The probability distribution of x_1 is determined completely, while $x_2, x_3 \cdots, x_n$ are still described by uncertain variables, and the uncertainty propagations turns into level-1 type. At stage 4, the knowledge of reliability analyst improves by obtaining the probability distributions type of x_2 and x_3, but their shaping parameters are still lacking and can be described by uncertain variables. x_1 and $x_4 \cdots, x_n$ are perfectly described by probability distributions. In this situation, the uncertainty propagations turn into level-2 type. At stage 5, the probability distributions of all input factors are determined completely due to elimination of epistemic uncertainties, and random variables are used to describe all input factors. Meanwhile, the uncertainty propagations turn into level-1 type.

In summary, the selection of uncertainty propagation types depends on the knowledge stage and sample data owned by the reliability analyst. For more general situations, the decision-makers can match any circumstances in practical engineering by increasing the number of stages.

5.2. Level-1 uncertainty propagation in structural reliability assessment

As the example mentioned in Section 5.1, there are three different knowledge stages in level-1 uncertainty propagation type, namely, stage 1, stage 3 and stage 5. For stage 1, the uncertain variables describe all the input factors of LSF, and uncertainty propagations analysis is handled through a pure uncertainty model. The uncertain reliability and index can be obtained based on the uncertainty theory and the operational laws, and the specific calculation methods of uncertain structure can be referred to the literature [35]. In stage 3, some input factors have sufficient data, so their probability distributions can be obtained, while some input factors lack data and can only be described by uncertain variables. By analyzing the joint propagation of uncertainty and probability, the methods proposed in the Sections 4.1 and 4.2 are used to estimate the reliability and index of uncertain random structure. The structure that corresponds to stage 5 is called random structure, and the uncertainty propagations analysis for random

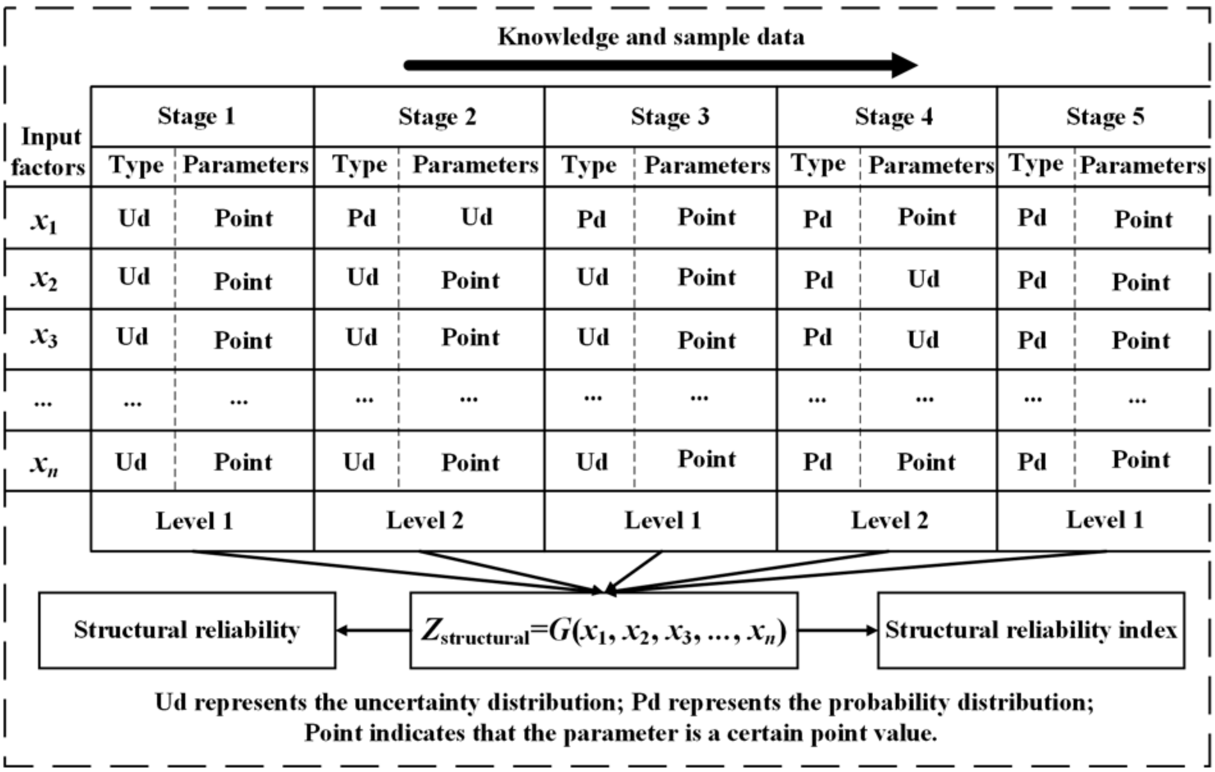


Fig. 1. Uncertainty propagation types of structure corresponding to different stages

structure can be implemented by traditional pure probability model. The probabilistic reliability and index of random structure can be estimated by the classical methods such as FORM and MCS.

5.3. Level-2 joint uncertainty propagation in uncertain random structure reliability assessment

As the example mentioned in Section 5.1, there are two different knowledge stages in the level-2 uncertainty propagation type, namely, stage 2 and stage 4. For stage 2, some input factors in the LSF are expressed by probability distributions, of which the shaping parameters are described by uncertainty distributions, while other input factors can be expressed by uncertainty distributions. The next two subsections will introduce the detailed calculation methods in these two stages.

Let a LSF of structural system contain $p+q$ input factors, of which p input factors are expressed by random variables, and the shaping parameters of the probability distributions are described by uncertain variables, while q input factors are expressed by uncertain variables. The level-2 joint propagation in stage 2 can be considered as a more general situation of propagations in stage 3. Consequently, a new level-2 uncertainty analysis method and the corresponding reliability calculation model are provided for stage 2 in this work.

Assuming that p input variables are represented as $\eta_1,\eta_2,\cdots,\eta_p$, and each probability distribution of η_i is represented as $\Psi_i(\eta_i|\theta_i)$, in which θ_i represents the shaping parameters of probability distribution. The shaping parameters are described by uncertainty distributions $\Phi_i(\theta_i)$. Let $\tau_1,\tau_2,\cdots,\tau_q$ represent the q input variables, and $\Upsilon_j(\tau_j)$ represent each uncertainty distribution of τ_j. According to the method presented in Section 4, if the LSF $G(\eta_1,\cdots,\eta_p;\ \tau_1,\cdots,\tau_q;\ \alpha)$ is continuous and strictly increasing with respect to $\tau_1,\tau_2,\cdots,\tau_k$, and strictly decreasing with respect to $\tau_{k+1},\tau_{k+2},\cdots,\tau_q$. Therefore, the chance reliability of the structural system at α can be calculated as:

$$\mathrm{Ch}_{\mathrm{reliability}}(\alpha)=\int_{\Re^p}F(0;y_1,y_2,\cdots,y_p;\alpha)\mathrm{d}\Psi_1(y_1|\theta_1)\mathrm{d}\Psi_2(y_2|\theta_2)\cdots\mathrm{d}\Psi_p(y_p|\theta_p) \tag{21}$$

where $F(0;y_1,y_2,\cdots,y_p;\alpha)$ is the root u of the Equation (15) for any real numbers $y_1,y_2,\cdots,y_p$.

Moreover, the CRI of the structural system at α can be computed from Equation (17), where the expected value and variance of LSF $G(\xi,\alpha)$ at α can be computed as follows:

$$E[G(\xi,\alpha)]=\int_{\Re^p}\int_0^1 G(y_1,\cdots,y_p;\Upsilon_1^{-1}(u),\cdots,\Upsilon_k^{-1}(u),\Upsilon_{k+1}^{-1}(1-u),\cdots,\Upsilon_q^{-1}(1-u);\alpha)\mathrm{d}u\mathrm{d}\Psi_1(y_1|\theta_1)\mathrm{d}\Psi_2(y_2|\theta_2)\cdots\mathrm{d}\Psi_p(y_p|\theta_p) \tag{22}$$

$$V[G(\xi,\alpha)]=2\int_{\Re^p}\int_0^{+\infty}x(1-F(e_{G(\xi,\alpha)}+x;y_1,\cdots,y_p;\alpha)+F(e_{G(\xi,\alpha)}-x;y_1,\cdots,y_p;\alpha))\mathrm{d}x\mathrm{d}\Psi_1(y_1|\theta_1)\mathrm{d}\Psi_2(y_2|\theta_2)\cdots\mathrm{d}\Psi_p(y_p|\theta_p) \tag{23}$$

where $e_{G(\xi,\alpha)}=E[G(\xi,\alpha)]$, and $F(x;y_1,\cdots,y_p;\alpha)$ is the root u of the Equation (20) for any real numbers $y_1,y_2,\cdots,y_p$.

Therefore, the chance reliability and CRI of the structural system is no longer a point value, but varies between the lower and upper bounds of shaping parameters with the uncertainty distribution $\Phi_i(\theta_i)$.

The level-2 joint propagation in stage 4 represents more general circumstances of random structure in stage 5. Thus, the reliability and index of random structure can also calculated by the traditional pure probability methods such as FORM and MCS. Consider a LSF $G(\eta_1,\cdots,\eta_n;\ \alpha)$ of structure contains n input factors, m input factors are expressed by the probability distributions $\Psi_i(\eta_i|\theta_i)$, and the shaping parameters θ_i of $\Psi_i(\eta_i|\theta_i)$ are described by the uncertainty distributions $\Phi_i(\theta_i)$, while $n-m$ input factors are expressed by random variables with no epistemic uncertainties. The variation range of reliability and index can be calculated by replacing the original probability distributions with $\Psi_i(\eta_i|\theta_i)$ in classical FORM. Thus, the probabilistic reliability and index of the structure are also no longer a point value, and the variation range can be obtained based on the uncertainty distribution of the shaping parameters.

6. An illustrated example

In this section, the propagation analysis methods developed herein, are applied to a practical engineering application of turbine disk reliability assessment. The description of turbine disk and the implementation of UD-LS surrogate model are introduced in Section 6.1. Section 6.2 shows the specific application process of the uncertainty propagation method proposed in this work. Some results and discussions on the advantages of the proposed method are given in Section 6.3.

Table 1. Input factors of three-stage turbine disk

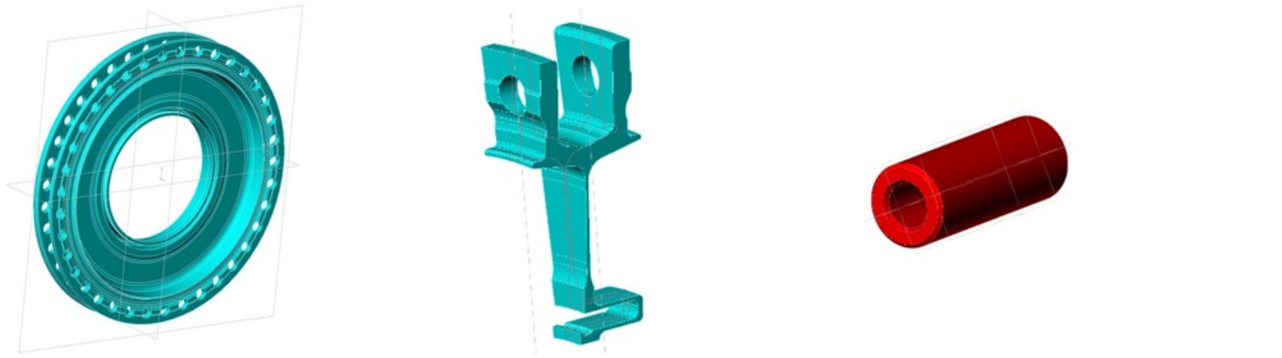
Input factors	Physical meaning	Mean value	Standard deviation
E_1(GPa)	Elastic modulus of roulette wheel	123	5
v_1	Poisson's ratio of roulette wheel	0.33	0.015
ρ_1(g/ cm^3)	Density of roulette wheel	4.48	0.2
E_2(GPa)	Elastic modulus of hollow pin shaft	219	10
v_2	Poisson's ratio of hollow pin shaft	0.3	0.015
ρ_2(g/ cm^3)	Density of hollow pin shaft	7.76	0.3
F(KN)	Resultant force on hollow pin shaft	24.925	0.315

6.1. Structure description and LSF simulation

Turbine disk is the key rotating component of modern aircraft engines, driven based on high-temperature gas in the engine combustion chamber. Because the turbine disk converts the thermal energy in the gas into mechanical energy to drive the engine, its reliability level will directly affect the performance of the entire engine.

As shown in Fig. 2(a), the three-stage turbine disk of a low-pressure compressor in a turbofan engine was selected as the research object. The pins on the roulette wheel are evenly and symmetrically distributed along the circumference. According to engineering analysis, the chief input factors affecting the reliability of roulette wheel include material characteristics, load and speed. In this work, the material of hollow pin was 3Cr13, while the material of roulette wheel was TC11. The blade load was applied perpendicularly to the hollow pin, and the average value of the load on each hollow pin was 24925N. The relevant parameters of input factors are shown in Table 1. Since the main failure mode of the turbine disk requires that the maximum stress value is greater than the allowable strength $S_{\mathrm{threshold}}$, the maximum stress value can be obtained by finite element analysis.

Because the shape and load of the turbine disk are completely symmetrical, 1/37 part of the turbine disk is considered to describe the entire structure. The average value of each input factor was chosen as the variable value, and the turbine disk was simulated using ANSYS 18.2 at a speed of 1000 rad/s. According to the simulation results



(a) Three dimensional model of the roulette wheel (b) 1/37 of the roulette wheel (c) Three dimensional model of hollow pin

Fig. 2. Model of three-stage turbine disk

Table 2. Finite element simulation results with uniform design

Sample number	$\omega(\mathrm{rad/s})$	$E_1(\mathrm{GPa})$	v_1	$\rho_1(\mathrm{g/cm^3})$	$E_2(\mathrm{GPa})$	v_2	$\rho_2(\mathrm{g/cm^3})$	F(KN)	$S_{\max}(\mathrm{MPa})$
1	0	110.50	0.3000	4.18	209.00	0.2925	7.910	25.240	631.411289
2	85	114.25	0.3187	4.53	231.50	0.3337	7.085	24.531	610.493275
3	170	118.00	0.3375	4.88	189.00	0.2775	8.210	25.870	635.223855
4	255	121.75	0.3562	3.93	211.50	0.3187	7.385	25.161	620.446302
5	340	125.50	0.3750	4.28	234.00	0.2625	8.510	24.453	653.177425
6	425	129.25	0.2962	4.63	191.50	0.3038	7.685	25.791	685.938023
7	510	133.00	0.3150	4.98	214.00	0.3450	6.860	25.082	684.244701
8	595	136.75	0.3337	4.03	236.50	0.2888	7.985	24.374	689.337371
9	680	108.00	0.3525	4.38	194.00	0.3300	7.160	25.713	739.236521
10	765	111.75	0.3712	4.73	216.50	0.2737	8.285	25.004	746.209250
11	850	115.50	0.2925	5.08	239.00	0.3150	7.460	24.295	776.036791
12	935	119.25	0.3112	4.13	196.50	0.2587	8.585	25.634	792.389327
13	1020	123.00	0.3300	4.48	219.00	0.3000	7.760	24.925	804.746581
14	1105	126.75	0.3488	4.83	241.50	0.3413	6.935	24.216	819.019637
15	1190	130.50	0.3675	3.88	199.00	0.2850	8.060	25.555	826.739102
16	1275	134.25	0.2887	4.23	221.50	0.3262	7.235	24.846	863.456674
17	1360	138.00	0.3075	4.58	244.00	0.2700	8.360	24.137	877.634253
18	1445	109.25	0.3262	4.93	201.50	0.3113	7.535	25.476	893.024436
19	1530	113.00	0.3450	3.98	224.00	0.2550	8.660	24.768	912.663743
20	1615	116.75	0.3637	4.33	246.50	0.2963	7.835	24.059	934.261452
21	1700	120.50	0.2850	4.68	204.00	0.3375	7.010	25.398	999.430433
22	1785	124.25	0.3037	5.03	226.50	0.2813	8.135	24.689	1027.33931
23	1870	128.00	0.3225	4.08	249.00	0.3225	7.310	23.980	981.648181
24	1955	131.75	0.3412	4.43	206.50	0.2662	8.435	25.319	1071.52881
25	2040	135.50	0.3600	4.78	229.00	0.3075	7.460	24.610	1114.01504

presented in Fig. 3, the stress-strain level at the junction between the roulette wheel and hollow pin is the highest, which is the dangerous failure point of the structure.

Since with the increase in rotating speed $\omega(\mathrm{rad/s})$, the maximum stress at the dangerous point will increase, the reliability of turbine disk will continue to degrade. According to the UD-LS surrogate model introduced in Section 3, a UD table $U_8^*(25^8)$ with 8 factors and 25 levels was designed to arrange the experiment. Let the speed range is $0 \sim 2040\,\mathrm{rad/s}$, and the range of other input factors is $x_i = \mu_i \pm f\sigma_i/4$, where μ_i and σ_i are the mean value and standard deviation of each input factors, respectively, and $f = 0,1,2,\cdots,12$. Then, finite element simulation can be used to calculate the maximum stress of each experiment. The simulation results corresponding to each experiment are shown in Table 2.

According to the stress-strength interference model, the LSF of the three-stage turbine disk is established as:

$$G(\boldsymbol{x}) = S_{\text{threshold}} - (b_0 + \sum_{i=1}^{8} b_i x_i + \sum_{j=9}^{16} b_j x_j^2) \tag{24}$$

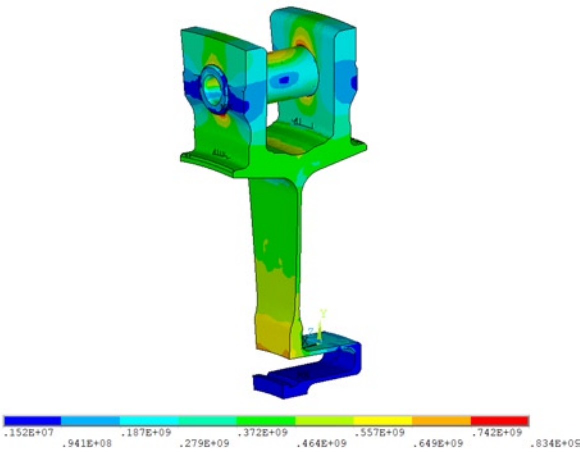


Fig. 3. The finite element stress cloud diagram of turbine disk at ω=1000 rad/s.

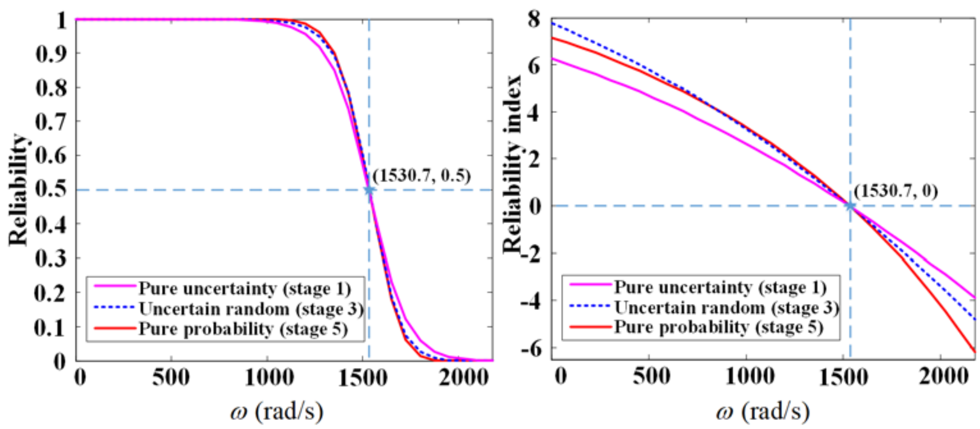
the LSF simulated in this experiment has a high degree of fit, which lays a good foundation for the next step of uncertainty propagation analysis.

6.2. Joint propagation of uncertainty and probability

Based on the different knowledge and sample data stages possessed by the reliability analyst on input factors, the reliability assessment can be implemented based on the uncertainty propagation model proposed in Section 5. According to the selection principles proposed in Section 5.1, the uncertainty propagation of the turbine disk can be obtained as shown in Table 3. In stage 1, the reliability analyst does not have detailed sample data on all input variables. So a domain expert is invited to estimate the values of input factors. Seven normal uncertainty distributions are used to represent the expert's beliefs corresponding to the input factors. In stage 2, the knowledge stage of reliability analyst is improved, and the distribution of roulette wheel density ρ_1 is confirmed as a normal probability distribution $N(\mu_{\rho_1},\sigma_{\rho_1})$. Nonetheless, the expected value μ_{ρ_1} of $N(\mu_{\rho_1},\sigma_{\rho_1})$ is still uncertain, and a domain expert is invited to estimate the values of μ_{ρ_1}. Therefore, a linear uncertainty distribution $\mathcal{L}(3.78,5.18)$ is used to represent the expert's beliefs on the expected value μ_{ρ_1}, but the other six input factors are still expressed as normal uncertainty distributions.

Table 3. Distribution types and parameters at different knowledge stages

Stages	Distribution types and parameters of input variables						
	$\rho_1(\mathrm{g/cm^3})$	$E_1(\mathrm{GPa})$	v_1	$E_2(\mathrm{GPa})$	v_2	$\rho_2(\mathrm{g/cm^3})$	$F(\mathrm{KN})$
Stage 1	$\mathcal{N}(4.48,0.2)$	$\mathcal{N}(123,5)$	$\mathcal{N}(0.33,0.015)$	$\mathcal{N}(219,10)$	$\mathcal{N}(0.3,0.015)$	$\mathcal{N}(7.76,0.3)$	$\mathcal{N}(24.925,0.315)$
Stage 2	$N(\mu_{\rho_1},0.2)$, $\mu_{\rho_1}\sim\mathcal{L}(3.78,5.18)$	$\mathcal{N}(123,5)$	$\mathcal{N}(0.33,0.015)$	$\mathcal{N}(219,10)$	$\mathcal{N}(0.3,0.015)$	$\mathcal{N}(7.76,0.3)$	$\mathcal{N}(24.925,0.315)$
Stage 3	$N(4.48,0.2)$	$\mathcal{N}(123,5)$	$\mathcal{N}(0.33,0.015)$	$\mathcal{N}(219,10)$	$\mathcal{N}(0.3,0.015)$	$\mathcal{N}(7.76,0.3)$	$\mathcal{N}(24.925,0.315)$
Stage 4	$N(4.48,0.2)$	$N(\mu_{E_1},5)$, $\mu_{E_1}\sim\mathcal{L}(95,151)$	$N(\mu_{v_1},0.015)$, $\mu_{v_1}\sim\mathcal{L}(0.274,0.386)$	$N(219,10)$	$N(0.3,0.015)$	$N(7.76,0.3)$	$N(24.925,0.315)$
Stage 5	$N(4.48,0.2)$	$N(123,5)$	$N(0.33,0.015)$	$N(219,10)$	$N(0.3,0.015)$	$N(7.76,0.3)$	$N(24.925,0.315)$



(a) Reliability of turbine disk in level-1 propagation

(b) Reliability index of turbine disk in level-1 propagation

Fig. 4. Reliability assessment in level-1 propagation

In stage 3, the knowledge stage of reliability analyst is improved further by obtaining sufficient data about roulette wheel density. So the normal probability distribution of ρ_1 is determined completely. Also, the normal uncertainty distributions of the other six input factors remain unchanged. In stage 4, the knowledge stage of reliability analyst is improved, and the distributions of seven input variables is determined as a normal probability distribution. However, the expected values of $N(\mu_{E_1},\sigma_{E_1})$ and $N(\mu_{v_1},\sigma_{v_1})$ are still unknown, and domain experts believe that the expected values of $N(\mu_{E_1},\sigma_{E_1})$ and $N(\mu_{v_1},\sigma_{v_1})$ obey the linear uncertainty distributions $\mathcal{L}(95,151)$ and $\mathcal{L}(0.274,0.386)$, respectively. In stage 5, the expected values of E_1 and v_1 are determined completely thanks to the sufficient sample data.

where $S_{\text{threshold}} = 935\text{Mpa}$ is the threshold of roulette wheel strength, $\boldsymbol{x} = x_1,\cdots x_8$ is a vector of eight input factors, and $\boldsymbol{b} = (b_0,b_1,\cdots,b_{16})^{\mathrm{T}}$ is the vector of coefficients, which is estimated by the method introduced in Section 3. The coefficient of determination is calculated as $R^2 = 0.99784$ by Equation (10), and is very close to 1. Hence, In other words, all input factors are perfectly described by a normal probability distribution. The above-mentioned specific reliability assessment processes under different knowledge and sample data stages are based on the joint uncertainty propagation method proposed in Section 5.

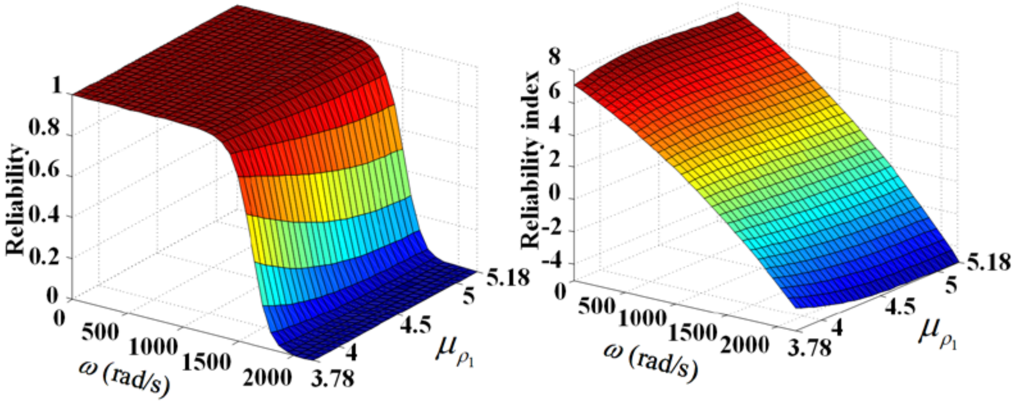
6.3. Results and discussion

Let the range of turbine disk speed ω be $[0, 2175]$. Then the reliability and index depending on ω in different stages can be calculated based on the methods developed in this paper. It is clear that with the increase in speed, the reliability and indexes of the turbine disk will degenerate because of the increase in stress.

As shown in Fig. 4, the reliability and indexes under three different stages in level-1, namely pure uncertainty in stage 1, uncertain random in stage 3 and pure probability in stage 5 are compared. The results estimated from level-2 in stage 2 are shown in Fig. 5, where the reliability and index of turbine disk fluctuate with the unknown parameter μ_{ρ_1}. In particular, when the rotation speed $\omega = 1200\text{rad/s}$, the reliability and index takes the values $(0.9648,\ 0.9895)$ and $(1.8533,\ 2.3081)$, respectively. Fig. 6 shows the variation of reliability and index with unknown parameters μ_{E_1} and μ_{v_1} at speed $\omega = 1200\text{rad/s}$, where the reliability and index take values in $(0.9445,\ 0.9975)$ and $(1.5935,\ 2.8018)$, respectively. The practical engineering example illustrates the specific implementation process of the presented method in detail, and the reliability of turbine disk are obtained in different knowledge and sample data stages.

Besides, the simulation results of the reliability and index for different knowledge stages at a specific speed $\omega = 1200\text{rad/s}$ are presented in Table 4. It is worth noting that the specific speed is selected arbitrarily and the same comparisons can be implemented at any speed value. As presented in Table 4, the reliability is transformed from the interval $(0.9648,\ 0.9895)$ in stage 2 to the determined value 0.9748 in stage 3 due to the increase in sample data. Moreover, the reliability is transformed from the interval $(0.9445,\ 0.9975)$ in stage 4 to the determined value 0.9871 in stage 5, which is a good explanation for the process of eliminating epistemic uncertainties. Similar conclusions can be obtained from the reliability indexes in different stages. Besides, as shown in Fig. 4 and Fig.5, when the reliability values are close to 1 at some speed values, the reliability index can be employed to distinguish the reliability differences at these speed values.

In the context of structural reliability assessment, the description of epistemic uncertainty is an inevitably common problem. Classical probability theory cannot be employed to express epistemic uncertainty since the real frequency cannot be obtained due to lack of data. Fuzzy measure and possibility measure do not satisfy duality property, and hence the description of epistemic uncertainty is not reasonable enough. Evidence and interval theory leads to the problem of interval expansion in practical applications. Uncertainty theory is a newly proposed mathematical framework that firmly conforms to the normality, duality and subadditivity theorems. This paper uses the uncertainty theory to describe epistemic uncertainty because it is more suitable for describing the human thinking processes. Also, the probability theory is chosen to represent aleatory uncertainty, and the chance theory is selected to deal with the situation when aleatory and epistemic uncertainties exist simultaneously. The results of case study shows that the level-1 and level-2 joint propagation can be explained very well by combining the above three theories. Consequently, the practical engineering application shows that the various knowledge stages outcome the different reliability levels, and the results highlight that the presented methods are effective and could deliver clear messages to decision-makers.



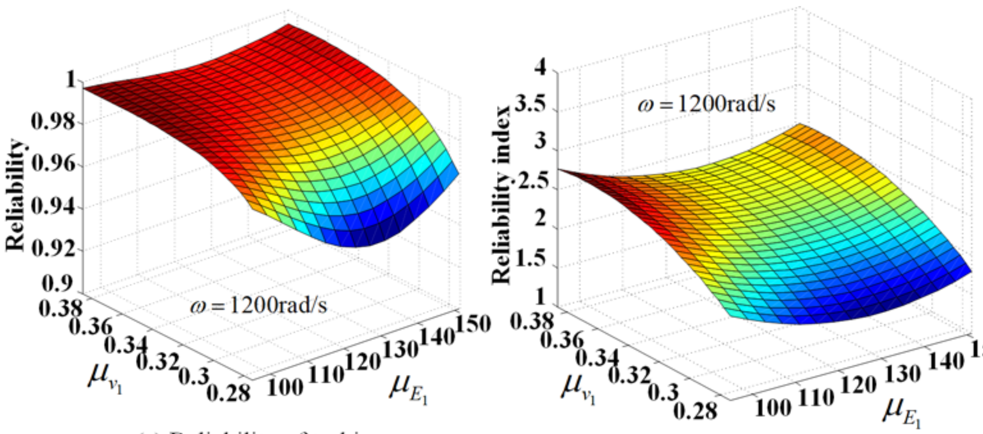
(a) Reliability of turbine disk in level-2 propagation

(b) Reliability index of turbine disk in level-2 propagation

Fig. 5. Reliability assessment in level-2 propagation (stage 2)

Table 4. Results of reliability and index corresponding to different stages

At ω=1200 rad/s	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Reliability	0.9565	(0.9648,0.9895)	0.9748	(0.9445,0.9975)	0.9871
Reliability index	1.7038	(1.8533,2.3081)	2.1092	(1.5935,2.8018)	2.2277



(a) Reliability of turbine disk in level-2 propagation

(b) Reliability index of turbine disk in level-2 propagation

Fig. 6. Reliability assessment of level-2 propagation at ω =1200 rad/s (stage 4)

7. Conclusions

In this paper, a novel uncertainty propagation method is proposed for the structural reliability assessment under mixed aleatory and epistemic uncertainties. To enable the analyst to calculate the structural reliability according to the different knowledge stages, the principles of selecting the uncertainty propagation types and the corresponding reliability estimation methods are presented. In summary, the main contributions of this paper are as follows:

(1) A new UD-LS surrogate model is proposed to solve the implicit LSF problem involving random and uncertain variables.

(2) The concepts of chance reliability and CRI are defined to describe structural reliability under mixed aleatory and epistemic uncertainties.

(3) A novel level-2 uncertainty analysis method and the corresponding reliability calculation model are provided for uncertain random structures.

(4) The principles for choosing reasonable uncertainty propagation types are presented for structural reliability assessment.

Decision-makers can evaluate the structural reliability corresponding to the different knowledge and sample data stages based on the uncertainty propagation method is proposed in this paper. As the model presented in this work is based on monotonic conditions, further research is required to focus on non-monotonic situations. Another interesting and important issue is to determine the distribution type of input factors based on small sample data.

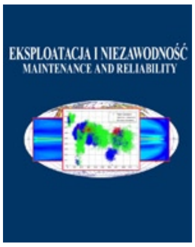
Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 51675026 and No. 71671009), and Aeronautical Science Foundation of China under Grant (2018ZC74001).

References

1. Cremona C, Gao Y. The possibilistic reliability theory: theoretical aspects and applications. Structural Safety 1997; 19(2): 173-201, https://doi.org/10.1016/S0167-4730(97)00093-3.
2. Durga Rao K, Kushwaha HS, Verma AK, et al. Quantification of epistemic and aleatory uncertainties in level-1 probabilistic safety assessment studies. Reliability Engineering & System Safety 2007; 92(7): 947-956, https://doi.org/10.1016/j.ress.2006.07.002.
3. Fang K. The uniform design: application of number-theoretic methods in experimental. Acta Mathematicae Applicatae Sinica 1980; 3(4): 363-372.
4. Fang P, Li S, Guo X, Wen Z. Response surface method based on uniform design and weighted least squares for non-probabilistic reliability analysis. International Journal for Numerical Methods in Engineering 2020; 121(18): 4050-4069, https://doi.org/10.1002/nme.6426.
5. Gao HY, Zhang XQ. Reliability-based design optimization under fuzzy and interval variables based on entropy theory. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(3): 430-439, https://doi.org/10.17531/ein.2019.3.9.
6. Guimaraes H, Matos JC, Henriques AA. An innovative adaptive sparse response surface method for structural reliability analysis. Structural Safety 2018; 73: 12-28, https://doi.org/10.1016/j.strusafe.2018.02.001.
7. Hu C, Youn BD, Wang PF. Engineering design under uncertainty and health prognostics. Springer 2019, https://doi.org/10.1007/978-3-319-92574-5.
8. Hu LH, Kang R, Pan X, et al. Risk assessment of uncertain random system-Level-1 and level-2 joint propagation of uncertainty and probability in fault tree analysis. Reliability Engineering & System Safety 2020; 198: 106874, https://doi.org/10.1016/j.ress.2020.106874.
9. Huang HZ. Structural reliability analysis using fuzzy sets theory. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14(4): 284-294.
10. Kiureghian AD, Ditlevsen O. Aleatory or epistemic? Does it matter? Structural Safety 2009; 31(2): 105-112, https://doi.org/10.1016/j.strusafe.2008.06.020.
11. Ke H, Liu H, Tian G. An uncertain random programming model for project scheduling problem. International Journal of Intelligent Systems 2015; 30(1): 66-79, https://doi.org/10.1002/int.21682.
12. Kang R, Zhang Q, Zeng Z, Zio E, Li X. Measuring reliability under epistemic uncertainty: review on non-probabilistic reliability metrics. Chinese Journal of Aeronautics 2016; 29(3): 571-579, https://doi.org/10.1016/j.cja.2016.04.004.
13. Li, HB, Nie, XB. Structural reliability analysis with fuzzy random variables using error principle. Engineering Applications of Artificial Intelligence 2018; 67: 91-99, https://doi.org/10.1016/j.engappai.2017.08.015.
14. Limbourg P, Rocquigny E, Andrianov G. Accelerated uncertainty propagation in two-level probabilistic studies under monotony. Reliability Engineering & System Safety 2010; 95:998-1010, https://doi.org/10.1016/j.ress.2010.04.012.
15. Liu B. Uncertainty Theory (5nd ed.). Berlin, Germany: Springer 2020.
16. Liu B. Why is there a need for uncertainty theory? Journal of Uncertain Systems 2012; 6(1): 3-10.
17. Liu B. Some research problems in uncertainty theory. Journal of Uncertain System 2009; 2(1): 3-10.
18. Liu B. Uncertainty Theory (2nd ed.). Berlin: Germany: Springer 2007.
19. Liu Y, Yao K. Option pricing formulas for uncertain exponential Ornstein-Uhlenbeck model with dividends. Soft Computing 2020; 20: 1-9, https://doi.org/10.1007/s00500-020-05177-z.
20. Liu YH. Uncertain random variables: a mixture of uncertainty and randomness. Soft Computing 2013; 17(4): 625-634, https://doi.org/10.1007/s00500-012-0935-0.
21. Liu YH. Uncertain random programming with applications. Fuzzy Optimization and Decision Making 2013; 12(2): 153-169, https://doi.org/10.1007/s10700-012-9149-2.
22. Malakzadeh K, Daei M. Hybrid FORM-Sampling simulation method for finding design point and importance vector in structural reliability. Applied Soft Computing 2020; 92: 1-13, https://doi.org/10.1016/j.asoc.2020.106313.
23. Melchers RE, Beck AT. Structural reliability analysis and prediction (3nd ed.). New York: John Wiley 2008.
24. Nirmala K, Babu KS, Reddy KH. Design and optimization of mechanical components and its mechanism using Monte Carlo simulation. International Journal of Engineering Science 2017; 3(6): 675-680.
25. Qin ZF. Uncertain random goal programming. Fuzzy Optimization and Decision Making 2018; 17(4): 375-386, https://doi.org/10.1007/s10700-017-9277-9.
26. Sohag K, Yiannis P. Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: A review, Safety Science 2019; 115: 154-175, https://doi.org/10.1016/j.ssci.2019.02.009.

27. Sun B, Li MM, Liao BP, et al. Time-variant reliability modeling based on hybrid non-probability method. Archive of Applied Mechanics 2020; 90(2): 209-219, https://doi.org/10.1007/s00419-019-01605-1.
28. Wang ZH, Wang ZL, Yu S. Time-dependent mechanism reliability analysis based on envelope function and vine-copula function. Mechanism and Machine Theory 2019; 134: 667-684, https://doi.org/10.1016/j.mechmachtheory.2019.01.008.
29. Wang W, Wang J, Fuj-h, Lu G-D. A moment-matching based method for the analysis of manipulator's repeatability of positioning with arbitrarily distributed joint clearances. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 10-20, https://doi.org/10.17531/ein.2019.1.2.
30. Wen M, Han Q, Yang Y, et al. Uncertain optimization model for multi-echelon spare parts supply system. Applied Soft Computing 2017; 56(C): 646-654, https://doi.org/10.1016/j.asoc.2016.07.057.
31. Wen M, Kang R. Reliability analysis in uncertain random system. Fuzzy Optimization and Decision Making 2016; 15(4): 491-506, https://doi.org/10.1007/s10700-016-9235-y.
32. Yao K, Zhou J, Ruin time of uncertain insurance risk process. IEEE Transactions on Fuzzy Systems 2018; 26(1): 19-28, https://doi.org/10.1109/TFUZZ.2016.2633329.
33. Zhang L, Zhang JG, Zhai H, Zhou S. A new assessment method of mechanism reliability based on chance measure under fuzzy and random uncertainties. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (2): 219-228, https://doi.org/10.17531/ein.2018.2.06.
34. Zhang Q, Kang R, Wen M. A new method of level-2 uncertainty analysis in risk assessment based on uncertainty theory. Soft Computing 2018; 22(17): 5867-5877, https://doi.org/10.1007/s00500-018-3337-0.
35. Zhou S, Zhang JG, Zhang L, et al. Advanced reliability analysis method for mechanisms based on uncertain measure. Journal of Intelligent & Fuzzy Systems 2020; 39(1): 1045-1059, https://doi.org/10.3233/JIFS-191970.
36. Zhang Q. A new approximate method for uncertainty propagation in system reliability analysis. Reliability Engineering & System Safety 1990; 29(2): 261-275, https://doi.org/10.1016/0951-8320(90)90081-W.
37. Zhang L, Zhang JG, You LF, Zhou S. Reliability analysis of structures based on a probability uncertainty hybrid model. Quality and Reliability Engineering International 2019; 35: 263-279, https://doi.org/10.1002/qre.2396.
38. Zhang Q, Kang R, Wen M. Belief reliability for uncertain random systems. IEEE Transactions on Fuzzy Systems 2018; 26(6): 3605-3614, https://doi.org/10.1109/TFUZZ.2018.2838560.
39. Zhang Z, Ruan XX, Duan MF, et al. An efficient epistemic uncertainty analysis method using evidence theory. Computer Methods in Applied Mechanics and Engineering 2018; 339: 443-466, https://doi.org/10.1016/j.cma.2018.04.033.
40. Zhao W, Fan F, Wang W. Non-linear partial least squares response surface method for structural reliability analysis. Reliability Engineering & System Safety 2017; 161: 69-77, https://doi.org/10.1016/j.ress.2017.01.004.



Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Gola A, Pastuszak Z, Relich M, Sobaszek Ł, Szwarc E. Scalability analysis of selected structures of a reconfigurable manufacturing system taking into account a reduction in machine tools reliability. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 242–252, http://doi.org/10.17531/ein.2021.2.4.

Indexed by: Web of Science Group

Scalability analysis of selected structures of a reconfigurable manufacturing system taking into account a reduction in machine tools reliability

Arkadiusz Gola[a], Zbigniew Pastuszak[b], Marcin Relich[c], Łukasz Sobaszek[a], Eryk Szwarc[d]

Faculty of Mechanical Engineering, Lublin University of Technology, ul. Nadbystrzycka 36, 20-618 Lublin, Poland
Faculty of Economics, Maria Curie Skłodowska University, Pl. Marii Curie-Skłodowskiej 5, 20-031 Lublin, Poland
Faculty of Economics and Management, University of Zielona Góra, ul. Podgórna 50d, 65-246 Zielona Góra, Poland
Faculty of Electronics and Computer Science, Koszalin University of Technology, ul. Śniadeckich 2, 75-453 Koszalin, Poland

Highlights

- The impact of a decline in the reliability of machine tools on the scalability of a RMS was assessed.
- · Two-, three-, four- and five-stage RMS structures were analyzed.
- · To identify bottlenecks and evaluate RMS productivity computer simulation methods were used.
- · The highest level of scalability was observed for the largest numbers of manufacturing stages.
- · The most stable level of reliability of the entire RMS was obtained for the lowest number of stages.

Abstract

Scalability is a key feature of reconfigurable manufacturing systems (RMS). It enables fast and cost-effective adaptation of their structure to sudden changes in product demand. In principle, it allows to adjust a system's production capacity to match the existing orders. However, scalability can also act as a "safety buffer" to ensure a required minimum level of productivity, even when there is a decline in the reliability of the machines that are part of the machine tool subsystem of a manufacturing system. In this article, we analysed selected functional structures of an RMS under design to see whether they could be expanded should the reliability of machine tools decrease making it impossible to achieve a defined level of productivity. We also investigated the impact of the expansion of the system on its reliability. To identify bottlenecks in the manufacturing process, we ran computer simulations in which the course of the manufacturing process was modelled and simulated for 2-, 3-, 4- and 5-stage RMS structures using Tecnomatix Plant Simulation software.

Keywords

reconfigurable manufacturing system, RMS, scalability, configuration, production structure, reliability, simulations, Tecnomatix Plant Simulation.

1. Introduction

At the end of the 20th century, manufacturing companies entered a new era, which, on the one hand, offered tremendous technical and IT solutions, but, on the other, brought them into competition with other firms not only on a local and national, but also on a global level [11]. To meet the requirements of the market, enterprises have to manufacture a wide range of products, constantly adjusting their product offerings to the changing demand. In order to maintain an appropriate level of competitiveness, companies must use manufacturing systems that allow to produce good quality commodities at a low production cost and quickly make the necessary changes to adapt to the incoming customer orders [14, 27]. These requirements can only be met by systems that combine the functional features of high-performance distributed manufacturing systems (DMS) and flexible manufacturing systems (FMS), and are, at the same time, scalable and dedicated to the processing of a particular family of products [3].

At the turn of the 21st century, a new concept of reconfigurable manufacturing systems (RMS) was developed to overcome the limitations of DMS and FMS. RMS, by definition, are designed for rapid change in structure that allows to adjust the system's functionality and production capacity to the current production requirements [22]. RMS, as a modern class of manufacturing systems, have an adaptive structure – both with regard to their hardware and software components, and are characterized by six core features: modularity, integrability, customized flexibility, diagnosability, convertibility, and scalability [7,39]. These characteristics provide a framework for the design of reconfigurable machine tools and reconfigurable controllers, the use of which allows to reduce the time-to-market and the costs of reconfiguring the manufacturing system [6].

Among these six characteristics, scalability is the one that is the most important from the point of view of the possibility of adapting a system to uncertain market changes by adjusting/ reconfiguring machines and/or the structure of the manufacturing processes [8]. Scalability is also a feature that allows to further optimize manufacturing systems and provides a basis for creating new manufacturing system paradigms focused on sustainable development and social welfare [32]. Moreover, scalability can be viewed as a buffer that allows to rapidly adjust a system's productivity in the event of a decrease in

E-mail addresses: A. Gola - a.gola@pollub.pl, Z. Pastuszak - z.pastuszak@umcs.lublin.pl, M. Relich - m.relich@wez.uz.zgora.pl, Ł. Sobaszek - l.sobaszek@pollub.pl, E. Szwarc - eryk.szwarc@tu.koszalin.pl

the reliability of its component machines and devices [9, 34].

Although from a technical and economic perspective, the scalability of a system should be defined at the stage of its design [23], this feature may actually be used for the analysis and optimization of the system's functioning throughout its service life [16]. As a rule, each manufacturing system is designed for a specific lifetime, which means that it should maintain an appropriate level of productivity throughout this period [31]. However, over time, the reliability of any technical system is bound to decrease [4, 21]. Given this, a system should have reserve production capacity as a buffer against a planned decrease in system reliability [38] or be designed so that the missing production capacity can be easily and cheaply offset [25, 29]. What is also important here are issues related to the nature and effectiveness of maintenance activities, which directly affect the dynamics of the decrease in a technical system's reliability level [10, 19].

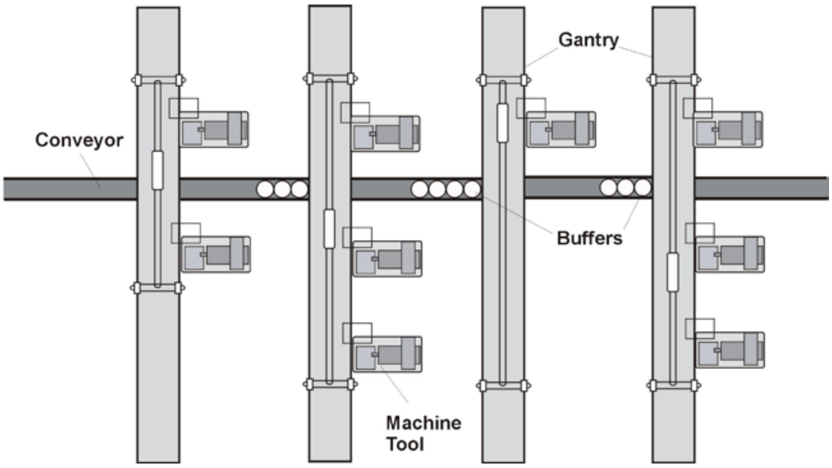


Fig. 1. Schematic of the arrangement of the structural components of an RMS (diagram prepared by the authors based on: [13])

In the case of RMS, a decline in reliability can be compensated for by adding new machine tools to the system to maintain an appropriate level of productivity [33]. Unfortunately, this type of solution entails costs related to both purchasing machine tools and securing appropriate production space. In this paper, we analyze the impact of reduced machine reliability on changes in the number of machines in a system and the system's reliability level. In particular, we examine selected system structures with different numbers of stages and different flexibility of machine tools. The study was conducted using computer simulation methods which are broadly applied in testing design assumptions in the processes of designing manufacturing systems, identifying bottlenecks, and increasing the efficiency of manufacturing systems (see, e.g. [20, 24]).

2. Scalability of RMS – a literature review

In designing manufacturing systems, designers focus on optimal selection of the systems' physical components, such as machine tools and means of in-plant transport, and their optimal arrangement, in order to meet pre-defined production requirements [2]. In the case of an RMS, these requirements may include optimization of the system's modular structure (which allows to reconfigure the system), optimal selection of the system's structure, and development of a system design that can accommodate changes in production demand.

A typical RMS consists of up to 20 stages with the machine tools of each stage having identical functional features (Fig. 1). In the machining process, parts are moved from one stage to the next using conveyors or overhead cranes. They are processed using CNC machine tools and/or RMS [8].

In order to adjust a system's production capacity (throughput) to changes in the needs of the market, the structure of the system must be reconfigured quickly and cost-effectively [1]. Production capacity is scaled in small and frequent discrete steps to smoothly adjust the system's functionality and throughput to match changes in customer demand [30]. As demonstrated by Putnik and colleagues [32], in practice, scalable capacity can be achieved by adding or removing specific machines, which is possible due to the parallel arrangement of the components of an RMS structure (Fig. 2).

Production capacity planning, understood as a problem of optimal adjustment of production capacity to the existing production needs and tasks, has been the subject of interest of numerous researchers in the last 40 years. The first studies on increasing systems' production capacity were carried out by Manne [28], and later elaborated on by Luss [26] and many other scholars (see e.g. [17, 41]). Their approach, however, was static. Considering that modern production systems have to deal with a rapidly changing and uncertain demand and that there are constant advancements in methods of designing manufactur-

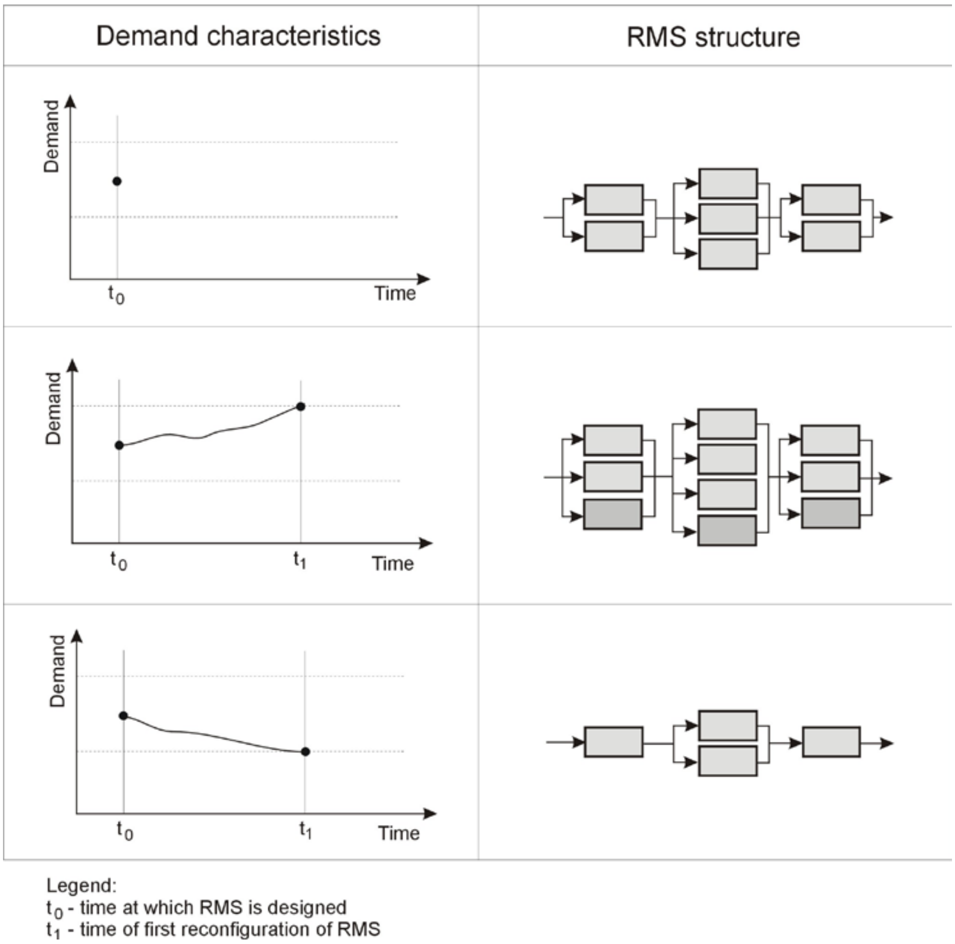


Fig. 2. Graphical illustration of RMS scalability based on [13]

ing systems, the problem of production capacity planning must be analyzed using a dynamic approach.

A review of the literature on scalability of manufacturing systems shows that there are two main lines of research in this area [8]:

1. design of RMS focused on increasing their scalability level, and
2. capacity planning using the scalability of RMS to adapt their production throughput to the existing demand.

Research on RMS design has been conducted, among others, by Spicer et al. [36], who investigated the problem of the impact of different system configurations (i.e. different degrees to which machines are arranged in parallel in a system's *n*-stage structures) on the system's productivity and production capacity. Son et al. studied the problems of stage paralleling and line balancing from the point of view of productivity and scalability of a production line. They showed that a completely balanced production line and an RMS could achieve an almost identical throughput, and that even an unbalanced RMS system generated smaller steps of production capacity changes than a balanced production line. Based on the results of a simulation study, Deif and ElMaraghy [12] proposed a new model for assessing system structures for different changing market demand scenarios. Wang and Koren [40] defined the scalability of a production system as its smallest possible incremental capacity change, and determined the relationship between the magnitude of this change and a system's scalability. Putnik et al. [32] conducted an extensive literature review in which he showed how_Wang and Koren's method could be used to assess *a posteriori* the scalability level of different configurations of RMS by comparing the throughput gain obtained for a specified number of additional machines or the cost needed to achieve a given level of productivity. Their conclusion was that a system's throughput and gain were higher in structures with a smaller number of stages. More recently, Hu et al. [18] analyzed the problem of joint optimization of production planning and adjustment of a system's production capacity based on product specifications, delivery time constraints and reconfigurable machine capabilities for assembly systems. Finally, Cerques et al. proposed their own metrics to evaluate the scalability of RMS by taking into account the parameters used for balancing operations on each stage of a production system [8].

The review of the literature shows that there are a large number of studies devoted to the problem of selecting an appropriate system structure in designing RMS. However, the focus of these studies is on the optimization of productivity and production capacity and their adaptation to the changing market demand. Unfortunately, the existing literature does not offer any empirical analyses of the impact of the decrease in the reliability of machine tools on the scalability of RMS over the system's entire service life, which is a large research gap.

To fill in this gap in research, we addressed the decision-making problem of choosing an appropriate RMS structure in experiments in which we evaluated selected RMS structures, taking into account the decrease in the level of reliability of machine tools during the system's service life. We used computer simulation methods and techniques for calculating the reliability of complex systems with hybrid structures, which permit to verify research assumptions without the need to build a physical model (a demonstrator). The goal of the study was to answer the question of how a decrease in the reliability of machine tools affects the need for expanding a scalable RMS and how it influences its reliability depending on the system's functional structure.

3. Research problem

In this study, which is a continuation of our earlier research reported in [15], we considered the problem of selection of the production structure of an RMS. As part of this study, we analyzed the structures of the RMS dedicated to the machining of body-type parts presented in article [22]. The decision-making problem under study can be formulated in the following way:

A *machine manufacturing company that provides machining services is planning to launch a new RMS production line for machining parts. The goal is to design an RMS dedicated to the machining of a body-type part (Fig. 4) which is produced in a technological process that encompasses five operations performed on two faces of the part, each face requiring separate fixturing (Fig. 4 b). The system under design should be capable of manufacturing a minimum of 500 parts a day. The working time per day for the RMS (F_j) is 1000 seconds.*

Over *time, the reliability of the individual machines decreases, which leads directly to a reduction in the system's productivity. If the existing system is not capable of producing 500 parts a day, it is expanded by adding another (new) machine tool at an appropriate location in the production line (a bottleneck). The main goal of this study was to find answers to the following questions.*

1) *How will the system be expanded (how many machine tools will be added in what locations) for each of the analyzed structures as the reliability of the machine tools decreases?*
2) *What level of reliability will the system achieve (for each structure) as it is scaled to the required productivity level?*

These questions need to be answered to identify the structures of the RMS under design that allow to maintain the assumed level of productivity while the level of reliability of machine tools decreases and new machines are added to the system.

4. Methods and results

As previous analyses and research findings for the analyzed RMS (see [15, 22]) show, the required productivity level of 500 parts a day can be achieved (assuming that all machine tools are 100% reliable) using one of the eight structures shown in Fig. 4. Because the production process must be carried out using at least two workholding fixtures (one for the execution of operation No. 10 and at least one for operations No. 20–50), we analyzed structures with from two (where operations No. 20–50 are executed using one type of multi-task ma-

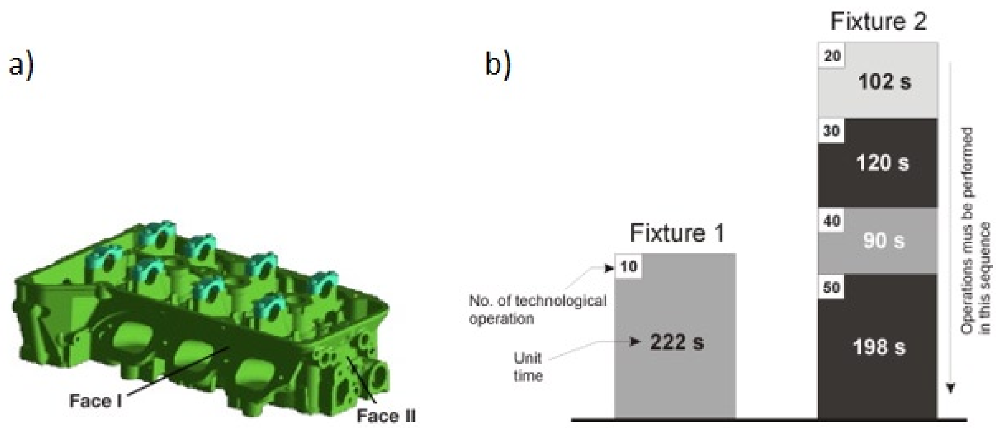


Fig. 3. Body-type part : a) general schematic view of the product, b) structure of the product's technological process

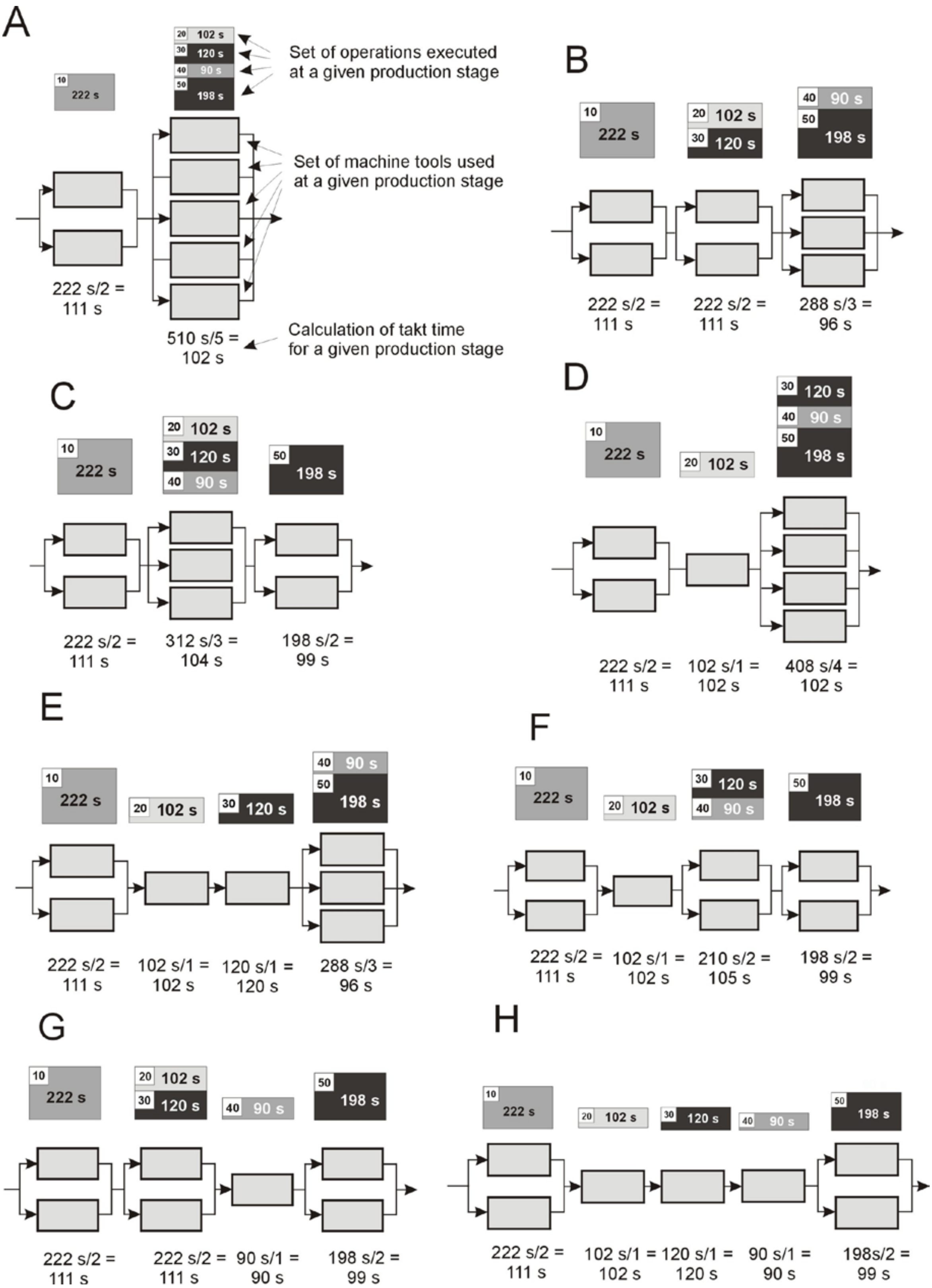


Fig. 4. RMS structures analyzed in this study

chines) to five stages (where each operation is performed on a different machine tool in the sequential stages of the process).

To answer the questions formulated in point 3, we carried out studies in which we:

– identified bottlenecks in the individual systems to find locations for system expansion in the event the reliability of the individual machine tools should decrease preventing the system from achieving the required productivity level;
– calculated the system's reliability level for each structure, taking into account the necessity of expanding the system to meet the existing production demand.

4.1. Analysis of the scalability of selected RMS structures as related to a decrease in system reliability

The scalability of an RMS, apart from allowing to dynamically adjust the system's structure to the existing production demand, also plays an important role as a "safety buffer" against wear and aging of the system's machines. Operation of any technical system is associated with a decrease in reliability, which translates into a reduction in its efficiency and productivity. In systems such as RMS, which ensure a short time-to-market and lower system expansion costs, the missing production capacity can be offset by adding new machines that will allow to execute the required production tasks. Obviously, excessive expansion of a system entails additional costs associated with purchasing machines and expanding the in-plant transport system as well as the need to find additional production space. This factor must be taken into account when selecting an appropriate system structure at the stage of designing an RMS.

We analyzed how a decline in the reliability of machine tools affected the expansion of the machine tool subsystem of the designed RMS for each of the eight structures shown in Fig. 4. To determine the impact of the decrease in reliability on the system's scalability, we assumed that the reliability of each machine tool was reduced by 1% in each observation period (this value was assumed to be sufficient to reliably interpret the results). Computer simulations were run to assess the impact of the decrease in reliability on the system's productivity. A Tecnomatix Plant Simulation model of the RMS was created for each of the eight structures, and a simulation of system operation was run, which covered a 1000-minute production period at a predefined level of reliability of the machine tools used in the production subsystem. An example of a model of the RMS developed for structure C (reliability level of 95%) is shown in Fig. 5.

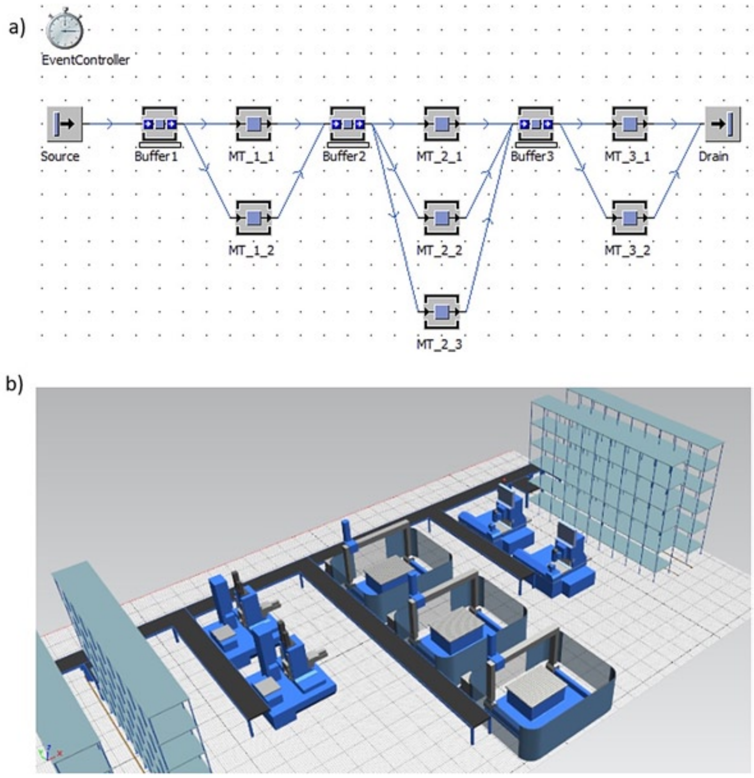


Fig. 5. A Tecnomatix Plant Simulation model for structure C: a) two-dimensional model of the machine tool subsystem, b) 3D visualization of the RMS

In the context of the design requirements defined earlier, the overriding goal is to maintain the system's production capacity at the level of minimum 500 parts per working day. When such a production volume cannot be obtained, it is necessary to identify the bottleneck (i.e. the production stage in which the machine tools have lost the ability to produce the specific number of products) and to eliminate it by "supplying" an additional machine tool that will provide reserve production capacity for the system's remaining service life. In this present study, it was assumed that each time the RMS's reliability is reduced, a new machine (with a 100% reliability level) with a functionality identical to that of the other machines at a particular stage of the system's structure is added to the system. For example, in the case of structure H, a drop in reliability of the base machines (machines that were originally in the system) to 93% makes it impossible to achieve a throughput of 500 parts (the system's productivity at this level of machine reliability is 495 pcs.). To compensate for this reduction, a new machine tool has to be added to stage I of the process, which is the bottleneck (Fig. 6). A general algorithm for the assessment of the impact of the decrease in machine reliability on system expansion is shown in Fig. 7.

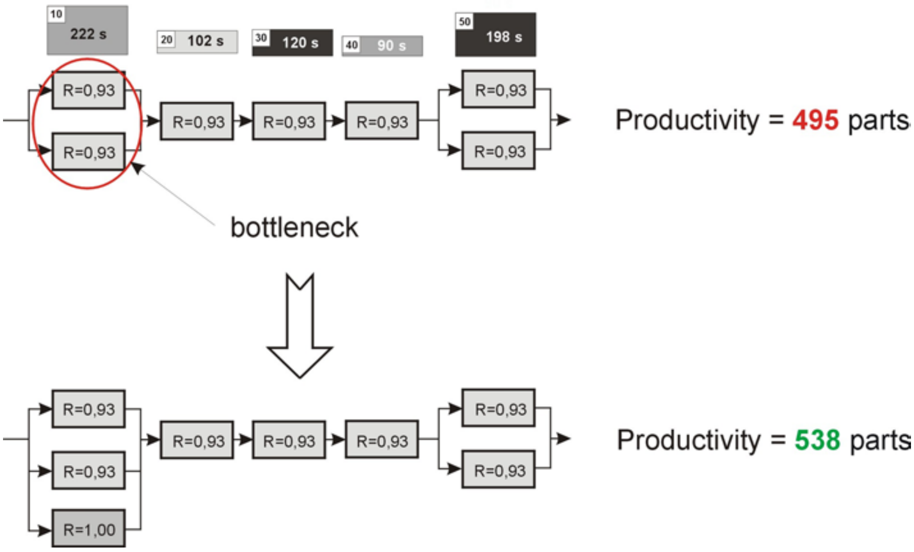


Fig. 6. Process of expanding the production structure of the RMS

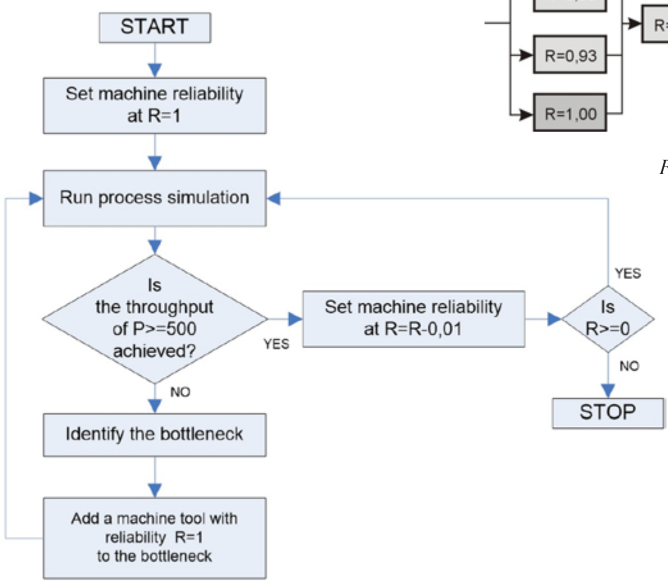


Fig. 7. Algorithm for scaling RMS in the event of reduction in machine tool reliability

Simulation experiments were carried out for each of the structures, in which the level of reliability of the base machines was reduced to from 99% to 1%. The results regarding the number of machine tools in each structure and the level of system productivity obtained are shown in Table 1 (to increase the transparency of the data, the tables show experimental and calculation results for every 5% decrease in reliability).

Table 1. System productivity and number of machine tools required to achieve the desired production target (with a division into production stages)

	Configuration							
R	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
1.00	2+5	2+2+3	2+3+2	2+4+1	2+1+2+3	2+1+2+2	2+2+1+2	2+1+2+1+2
	534	534	534	534	534	534	534	534
0.95	2+5	2+2+3	2+3+2	2+1+4	2+1+2+3	2+1+2+2	2+2+1+2	2+1+2+1+2
	509	508	509	506	508	508	506	506
0.90	3+5	3+3+3	3+3+2	3+1+4	3+1+2+3	3+1+2+2	3+3+1+2	3+1+2+1+2
	525	561	514	523	522	509	544	518
0.85	3+6	3+3+3	3+4+2	3+2+5	3+2+2+3	3+2+3+2	3+3+1+2	3+2+2+1+2
	616	531	514	646	531	514	514	514
0.80	3+6	3+3+4	3+4+3	3+2+5	3+2+2+4	3+2+3+3	3+3+1+3	3+2+2+1+3
	581	656	635	607	654	660	533	533
0.75	3+6	3+3+4	3+4+3	3+2+5	3+2+2+4	3+2+3+3	3+3+2+3	3+2+2+2+3
	545	614	600	572	615	619	621	617
0.70	3+6	3+3+4	3+4+3	3+2+5	3+2+2+4	3+2+3+3	3+3+2+3	3+2+2+2+3
	508	565	555	530	561	574	578	564
0.65	3+7	3+3+4	3+4+3	3+2+6	3+2+2+4	3+2+3+3	3+3+2+3	3+2+2+2+3
	540	533	522	541	526	531	532	521
0.60	4+7	4+3+4	4+5+3	4+2+6	4+2+2+4	4+2+3+3	4+4+2+3	4+2+2+2+3
	546	502	98	601	541	541	598	588
0.55	4+7	4+4+5	4+5+3	4+2+6	4+2+2+5	4+2+2+3	4+4+2+3	4+2+2+2+3
	504	697	553	552	549	503	553	540
0.50	4+8	4+4+5	4+5+3	4+2+6	4+2+3+5	4+2+4+3	4+4+2+3	4+2+3+2+4
	574	647	508	512	633	508	508	641
0.45	4+8	4+4+5	4+5+4	4+2+7	4+2+3+5	4+2+4+4	4+4+2+4	4+2+3+2+4
	527	594	526	600	574	601	599	578
0.40	4+9	4+4+5	4+6+4	4+2+7	4+2+3+5	4+2+4+4	4+4+2+4	4+2+3+2+4
	553	541	555	533	533	543	541	520
0.35	5+9	5+4+6	5+6+4	5+3+7	5+3+3+6	5+3+4+4	5+5+2+4	5+3+3+2+4
	546	501	607	506	697	563	626	624
0.30	5+10	5+5+6	5+6+4	5+3+8	5+3+3+6	5+3+4+4	5+5+2+4	5+3+3+2+4
	609	642	547	597	642	505	560	559
0.25	5+10	5+5+6	5+7+4	5+3+8	5+3+3+6	5+3+5+4	5+5+3+4	5+3+3+3+4
	551	582	510	536	577	513	510	507
0.20	5+11	5+5+6	5+7+5	5+3+9	5+3+3+6	5+3+5+5	5+5+3+5	5+3+3+3+5
	561	511	563	564	511	563	559	533
0.15	6+11	6+5+7	6+7+5	6+3+9	6+4+4+7	6+3+5+5	6+6+4+5	6+3+4+3+5
	544	504	537	560	661	577	663	655
0.10	6+12	6+6+7	6+8+5	6+3+10	6+4+4+7	6+3+5+5	6+6+4+5	6+3+4+3+5
	592	575	590	620	586	506	585	573
0.05	6+12	6+6+7	6+8+5	6+3+10	6+4+4+7	6+3+6+5	6+6+4+5	6+3+4+3+6
	527	503	516	538	514	510	511	519

Legend:

2+2+3	- system configuration (number of machine tools in each stage of the process)
508	- system productivity (number of products manufactured in 1000 minutes)

An analysis of the data given in Figure 8 shows that the largest number of machine tools had to be added to structures with the largest number of stages (configurations E, F, G, and H) to ensure the required production capacity level. Regardless of the level of decrease in machine reliability, in all cases, the smallest number of machine tools were added to the system with the smallest number of stages (structure A).

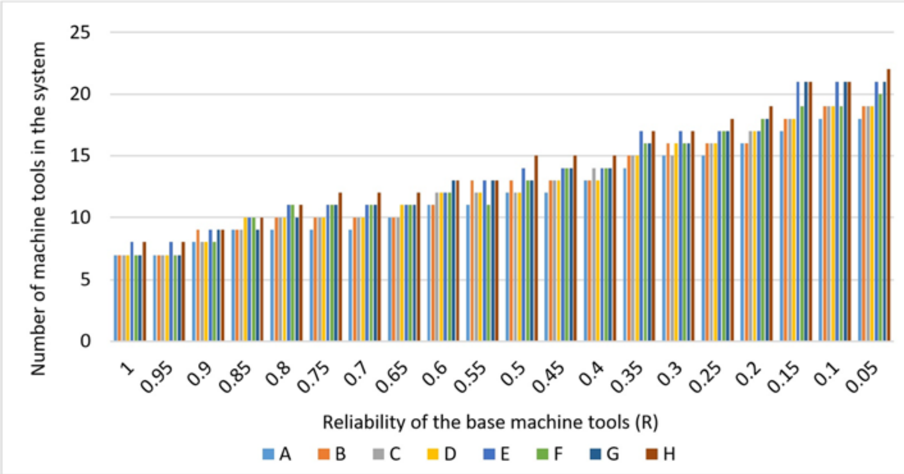


Fig. 8. Relationship between the number of machine tools in the system and reliability of the base machine tools

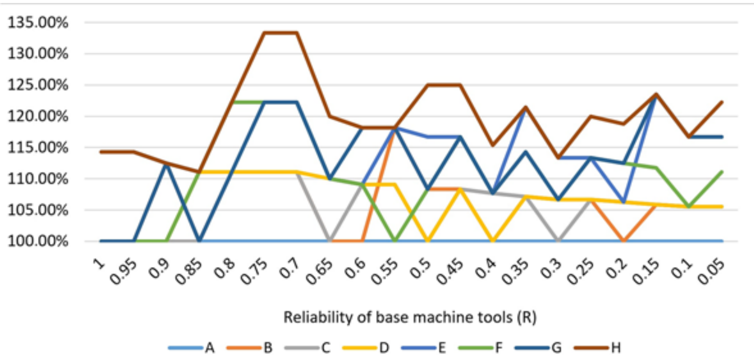


Fig. 9. A graph showing the percent increase in the number of machine tools in the individual RMS structures relative to structure A

The largest percent difference in the number of machine tools in relation to structure A was observed for structure H, which had the largest number of stages (Table 2, Fig. 9). The average percent increase in the number of machines relative to structure A ranged from 5.21% to 6.27% for the three-stage structures (B, C, and D), and from 5.35% to 12.32 % for the four-stage structures (E, F and G). The system with structure H (a five-stage structure) used 19.82% more machine tools than the system with two stages (the largest difference of 33.33% was found for machine tool reliability level of 70–75% (Table 2).

An important factor that needs to be considered in assessing RMS structures is the impact of scalability of a system on its productivity. In the case under study, the system is expanded by adding a new machine tool at a location identified as a bottleneck when the decrease in machine reliability makes it impossible to achieve the productivity level of 500 items per 1000 min (a day). Expansion of a system allows to maintain a required level of production capacity and, in many cases, also to build up production reserves as a buffer against a further decrease in productivity resulting from the aging of machines.

A system's scalability, in accordance with the principles of RMS, permits to dynamically adjust production capacity to the current production demand. To evaluate the impact of the investigated system's scalability on its productivity, simulation experiments were carried out for each structure in accordance with the algorithm presented in Fig. 5. The results are given in Figure 10

As shown in Figure 10, the system's productivity (production capacity) increases stepwise as new machine tools are added to the system. However, it increases slightly differently for each of the structures. When the increase in production capacity is

Table 2. Percent increase in the number of machine tools in the individual RMS structures relative to structure A

R	B	C	D	E	F	G	H
1	0.00%	0.00%	0.00%	14.29%	0.00%	0.00%	14.29%
0.95	0.00%	0.00%	0.00%	14.29%	0.00%	0.00%	14.29%
0.9	12.50%	0.00%	0.00%	12.50%	0.00%	12.50%	12.50%
0.85	0.00%	0.00%	11.11%	11.11%	11.11%	0.00%	11.11%
0.8	11.11%	11.11%	11.11%	22.22%	22.22%	11.11%	22.22%
0.75	11.11%	11.11%	11.11%	22.22%	22.22%	22.22%	33.33%
0.7	11.11%	11.11%	11.11%	22.22%	22.22%	22.22%	33.33%
0.65	0.00%	0.00%	10.00%	10.00%	10.00%	10.00%	20.00%
0.6	0.00%	9.09%	9.09%	9.09%	9.09%	18.18%	18.18%
0.55	18.18%	9.09%	9.09%	18.18%	0.00%	18.18%	18.18%
0.5	8.33%	0.00%	0.00%	16.67%	8.33%	8.33%	25.00%
0.45	8.33%	8.33%	8.33%	16.67%	16.67%	16.67%	25.00%
0.4	0.00%	7.69%	0.00%	7.69%	7.69%	7.69%	15.38%
0.35	7.14%	7.14%	7.14%	21.43%	14.29%	14.29%	21.43%
0.3	6.67%	0.00%	6.67%	13.33%	6.67%	6.67%	13.33%
0.25	6.67%	6.67%	6.67%	13.33%	13.33%	13.33%	20.00%
0.2	0.00%	6.25%	6.25%	6.25%	12.50%	12.50%	18.75%
0.15	5.88%	5.88%	5.88%	23.53%	11.76%	23.53%	23.53%
0.1	5.56%	5.56%	5.56%	16.67%	5.56%	16.67%	16.67%
0.05	5.56%	5.56%	5.56%	16.67%	11.11%	16.67%	22.22%
Mean	5.93%	5.21%	6.27%	15.35%	10.19%	12.32%	19.82%

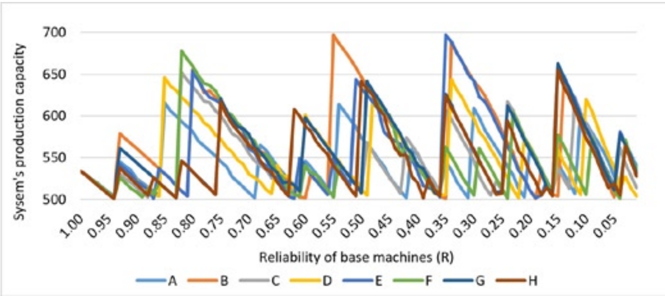


Fig. 10. Productivity of the RMS as a finction of the reliability of the base machine tools (taking into account system scalability)

considered over the range from 100% (R = 1.00) to 1% (R = 0.01) machine reliability, the smallest "leaps" in production capacity are observed for structure A (from 502 to 616 pcs/1000 min of system operation). The largest spread is observed for structures B and E (from 501 to 697 pcs/1000 min.). The mean productivity values for the analyzed scenarios are as follows: structure A: 543 pcs/1000 min, Structure B: 568 pcs/1000 min, Structure C: 549 pcs/1000 min, Structure D: 552 pcs/1000 min, Structure E: 562 pcs/1000 min, Structure F: 555 pcs/1000 min, Structure G: 549 pcs. A graphic interpretation of the statistical analysis of the results is shown in Fig. 11.

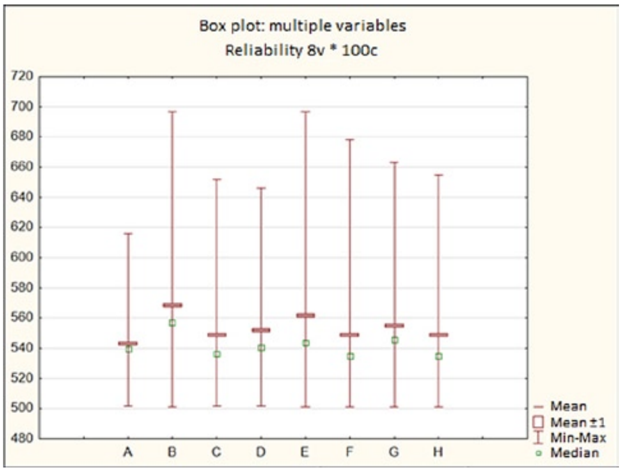


Fig. 11. Results of the statistical analysis of system productivity for the investigated RMS structures

The smallest spread between the minimum and the maximum increase in production capacity was observed for the two-stage structure (A). Taking into account the fact that smooth adjustment of production capacity to current demand is one of the key principles of RMS, structure A seems to be the most desirable. The largest "leaps" occurred in structures B and E. In practical terms, this means that these structures had excessive reserve capacity, which is not beneficial from the point of view of the economics of maintanence of technical equipment.

4.2. Evaluation of RMS reliability

The scalability of a system has a direct impact on its level of reliability. In the case under study, on the one hand, reliability constantly decreases as a consequence of the decline in the reliability of the individual machines, and on the other hand, it increases as new, 100% reliable machine tools are added to the system's structure. Considering that the total reliability of the system depends both on the reliability of all its components and the way they are arranged, the impact of expansion of the system on its reliability is an important factor that should be evaluated at the stage of selecting the appropriate functional structure of the designed RMS.

The reliability of a system is a derivative of both the number of production stages and the number of machines used in each stage. If components are added serially, the system's reliability is reduced. In a case like this, if the reliability of each machine is R, and the number of machines is n, then the reliability of the system is R^n. Parallel arrangement of two identical components increases the overall reliability of the system. More components added in parallel (Fig. 12 b) increase the reliability of the system, because the system will stop functioning only when all system components have failed. In this case, the probability that n identical machines arranged in parallel will fail is $(1–R)^n$, and the system's reliability is $1\text{-}(1\text{-}R)^n$ [37]. All of the RMS structures analyzed in this example are hybrids that combine the characteristics of both parallel and serial structures.

Calculations of the system's reliability for three selected structures are given in Table 3. To show precisely how the system's reliability was calculated, structures with different numbers of stages and different levels of reliability of the individual machine tools were selected.

The system's reliability level for each of the structures was calculated under the assumptions regarding the decrease in the reliability level of machine tools and system scalability presented in section 4.1 of this paper. The results of the calculations made for every 5% decrease in reliability are given in Table 4, and a graphic interpretation of the results is shown in Fig. 12.

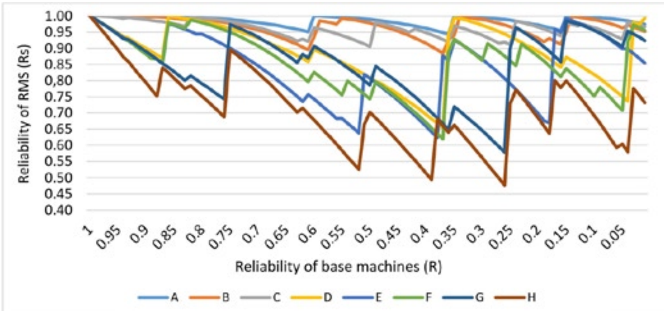


Fig. 12. Reliability of the scalable RMS (R_s) as a function of the decrease in the reliability of its base machines (R)

The highest mean level of system reliability of nearly 98.92% was observed for the RMS with a two-stage structure (structure A). The poorest result was obtained for the five-stage structure (structure H), for which the mean level of system reliability was only 72.65%. It is worth emphasizing that it was only systems with two- or three-stage structures that had an over 90% reliability, and the system's reliability decreased along with the increase in the number of processing stages (despite the fact that new machines characterized by 100% reliability were consistently added to the system).

The reliability curves for the analyzed period clearly show that the system with structure A had the highest and most stable level of reliability, while structures D, E, F, G and H were characterized by the largest leaps in reliability. This is confirmed by the summary_results of statistical analysis shown in Figure 13.

A detailed analysis of the results clearly shows that the two-stage structure (structure A) has the best properties from the point of view of system reliability over the entire period analyzed. The scalability of the system with this structure, despite the decrease in the reliability of the individual machines (from 99% to 1%), allows to maintain system reliability at the level from 94.3824% to 99.9998%. In the case of the three-stage structures (B and C), the system's reliability ranges from 88.5015% to 99.9999%, and for the remaining structures, it ranges from 47.6314% to 99.9799%. Considering the fact that the reliability of a system, in practice, translates into its flawless operation over the entire service life, this factor is key in selecting the appropriate system structure.

Table 3. Method of calculating the reliability level of the RMS under design

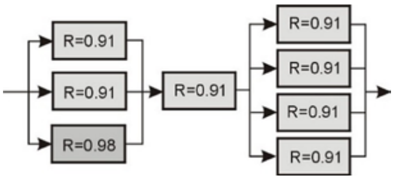
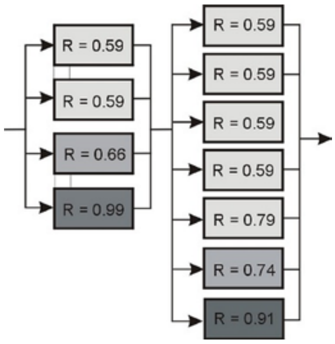
RMS structure	System reliability (R_S)
Structure H, reliability of base machines R = 0.94 	$R_S = [1 - (1 - 0.94)^2] * 0.94 * [1-(1-0.94)^2]* 0.94 * [1-(1-0.94)^2] =$ **0.874091**
Structure D, reliability of base machines R = 0.91 	$R_S = [1 - (1 - 0.91)^2 * (1 - 0.98)] * 0.91 * [1-(1-0.94)^4] =$ **0.909793**
Structure A, reliability of base machines R = 0.59 	$R_S = [1 - (1 - 0.59)^2 * (1-0.66) * (1-0.99)] * [1 - (1 - 0.59)^5 * (1-0.74) * (1-0.91)] =$ **0.999157**

Table 4. System reliability level for each of the structures of the scalable RMS

R	A	B	C	D	E	F	G	H
1	100.0000%	100.0000%	100.0000%	100.0000%	100.0000%	100.0000%	100.0000%	100.0000%
0.95	99.7500%	99.4882%	99.4882%	94.7619%	94.5138%	94.2893%	94.2893%	89.5748%
0.9	99.9690%	99.8401%	98.8713%	89.9640%	88.9842%	88.1825%	89.0465%	79.3643%
0.85	99.8200%	99.3040%	97.5675%	99.6703%	97.0989%	97.3619%	82.7887%	80.8286%
0.8	99.4784%	98.9627%	99.3448%	98.2784%	94.3548%	97.8933%	79.1068%	75.2708%
0.75	98.8653%	97.6863%	98.2578%	96.1184%	90.0758%	94.9574%	97.3349%	89.5211%
0.7	97.8943%	95.6439%	96.4277%	93.1161%	84.6097%	90.7285%	93.4441%	82.4031%
0.65	96.5670%	92.6579%	93.6670%	89.4721%	78.0068%	85.1405%	88.1195%	73.9058%
0.6	99.9795%	93.5076%	96.4467%	89.5713%	74.2614%	82.5733%	90.5954%	67.8813%
0.55	99.5436%	99.3086%	93.9766%	85.6135%	68.3080%	75.6173%	85.2320%	58.4092%
0.5	98.9172%	97.7842%	90.5059%	80.8523%	80.6762%	74.3454%	78.6073%	70.2166%
0.45	97.7812%	95.2754%	96.2272%	75.7465%	73.8451%	74.5410%	79.3085%	61.1457%
0.4	96.1827%	91.4958%	94.8327%	69.5560%	65.8819%	67.0279%	71.8052%	51.4000%
0.35	99.9782%	94.0611%	96.9576%	99.3216%	92.6218%	92.7751%	71.9428%	66.2636%
0.3	99.5370%	98.8758%	94.0071%	97.1635%	86.6447%	86.1818%	64.5615%	56.1794%
0.25	98.6472%	96.6963%	91.0543%	93.5488%	79.0229%	86.1307%	89.7804%	72.7402%
0.2	97.1968%	93.0106%	96.5654%	88.9814%	69.7254%	87.4975%	91.8359%	68.4206%
0.15	99.9875%	95.1974%	97.9106%	87.2665%	99.1206%	83.6950%	98.5187%	79.8808%
0.1	99.5791%	98.7923%	96.3801%	81.8705%	95.5585%	75.2534%	95.5665%	69.1655%
0.05	98.6743%	96.3186%	92.7535%	75.1184%	89.8394%	70.7615%	90.4742%	60.3720%
Mean	**98.9174%**	**96.6953%**	**96.0621%**	**89.2996%**	**85.1575%**	**85.2477%**	**86.6179%**	**72.6472%**

5. Conclusions and further research

In the process of designing a manufacturing system, it is necessary to consider aspects related to the system's entire service life. Particularly important in this respect is the problem of wear of machine tools and other components of the system, which reduces its reliability and, consequently, also its efficiency and productivity. For that reason, reliability issues should be analyzed already at the stage of creating a technical design.

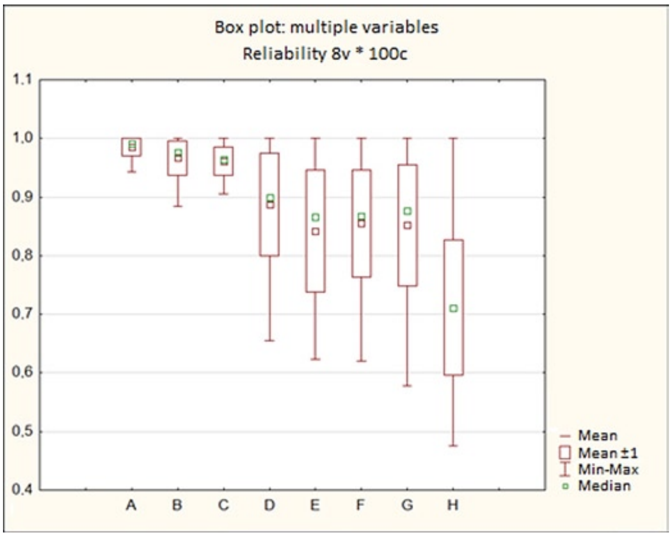


Fig. 13. Results of the statistical analysis of the reliability of the RMS for each of the analyzed structures

In the case of RMS, one of the key features of a structure being designed is scalability, which allows to adjust the system's production capacity to the existing market demand. This feature also permits to supplement a system's production capacity reduced by the decline in the reliability of the component machine tools, as discussed in the present study.

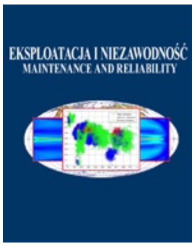
As part of the present experiments, we analyzed eight structures of an RMS dedicated to the production of body-type parts. In particular, we wanted to find answers to the following questions: (1) How will the system be expanded, for each structure, in order to ensure the minimum required level of system productivity? (2) How will system expansion contribute to building up production reserves for the production subsystem? (3) How will the level of reliability change over the system's service life? Computer simulation methods were used to evaluate the system's productivity and to identify bottlenecks. The system's operation was modelled and simulated for each of the eight structures, assuming that machine reliability decreased in a linear manner over the system's service life.

The results clearly indicate that RMS structures that have the best properties are those with the smallest number of stages. Systems with this type of structures, when expanded, show small increments in production capacity (and thus a minimum redundancy of production reserves) and exhibit the highest levels of reliability. Unfortunately, in practice, the use of structures with fewer stages requires the deployment of multi-task machine tools, generating higher per-unit purchase costs. Given all this, in our future research, we plan to carry out a multicriteria analysis, in which, apart from the functional and efficiency-related features of the individual structures, we will look into the economic aspects of system construction, such as the price of machine tools, the methods and costs of organizing a system's transport and storage subsystems, as well as the use of elements for controlling the individual components of a system in accordance with the assumptions of Industry 4.0.

References

1. Aleš Z, Pavlů J, Legát V, Mošna F, Jurča V. Methodology of overall equipment effectiveness calculation in the context of Industry 4.0 environment. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (3): 411-418, https://doi.org/10.17531/ein.2019.3.7.
2. AlGeddawy T, ElMaraghy HA. Changeability Effect on Manufacturing Systems Design [in:] ElMaraghy H. (eds), Changeable and Reconfigurable Manufacturing Systems. Springer Series in Advanced Manufacturing. Springer, London, 2009, https://doi.org/10.1007/978-1-84882-067-8_15.
3. Andersen A-L, Larsen JK, Brunoe TD, Nielsen K, Ketelsen C. Critical enablers of changeable and reconfigurable manufacturing and their industrial implementation. Journal of Manufacturing Technology Management 2018; 29(6): 983-1002, https://doi.org/10.1108/JMTM-04-2017-0073.
4. Antosz K. Maintenance - identification and analysis of the competency gap. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (3): 484-494, https://doi.org/10.17531/ein.2018.3.19.
5. Asl FM, Ulsoy G. Stochastic Optimal Capacity Management in Reconfigurable Manufacturing Systems. CIRP Annals 2003; 52(1): 371-374, https://doi.org/10.1016/S0007-8506(07)60604-0.
6. Battaïa O, Benyoucef L, Delorme X, Dolgui A, Thevenin S. Sustainable and Energy Efficient Reconfigurable Manufacturing Systems [in:] Benyoucef L. (ed.), Reconfigurable Manufacturing Systems: From Design to Implementation. Springer Series in Advanced Manufacturing. Springer, Cham, 2020, https://doi.org/10.1007/978-3-030-28782-5_9.
7. Bortolini M, Galizia FG, Mora C. Reconfigurable Manufacturing Systems: literature review and research trend. Journal of Manufacturing Systems 2018; 49: 93-106, https://doi.org/10.1016/j.jmsy.2018.09.005.
8. Cerqueus A, Delorme X, Dolgui A. Analysis of the Scalability for Different Configurations of Lines [in:] Benyoucef L. (ed.), Reconfigurable Manufacturing Systems: From Design to Implementation. Springer Series in Advanced Manufacturing. Springer, Cham, 2020, https://doi.org/10.1007/978-3-030-28782-5_7.
9. Dahane M, Menyoucef L. An adapted NSGA-II algorithm for a Reconfigurable manufacturing system (RMS) design under machines reliability constraints. Metaheuristics for Production Systems 60: 109-130, https://doi.org/10.1007/978-3-319-23350-5_5.
10. Daniewski K, Kosicka E, Mazurkiewicz D. Analysis of the Correctness of Determination of the Effectiveness of Maintenance Service Actions. Management and Production Engineering Review 2018; 9 (2): 20-25, https://doi.org/10.24425/119522.
11. Danilczuk W, Gola A., Computer-Aided Material Demand Planning Using ERP Systems and Business Intelligence Technology, Applied Computer Science 2020; 16(3): 42-55, https://doi.org/10.23743/acs-2020-20.
12. Deif AM, ElMaraghy HA. Assessing capacity scalability policies in RMS using system dynamics. International Journal of Flexible Manufacturing Systems 2007; 19(3): 128-150, https://doi.org/10.1007/s10696-008-9031-2.
13. Deif AM, ElMaraghy WH. A control approach to explore the dynamics of capacity scalability in reconfigurable manufacturing systems. Journal of Manufacturing Systems 2006; 25(1): 12-24, https://doi.org/10.1016/S0278-6125(07)00003-9.
14. Esmaeilian B, Behdad S, Wang B. The evolution and future of manufacturing: A review. Journal of Manufacturing Systems 2016; 39, 79-100, https://doi.org/10.1016/j.jmsy.2016.03.001.
15. Gola A. Reliability analysis of reconfigurable manufacturing structures using computer simulation methods. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 90-102, https://doi.org/10.17531/ein.2019.1.11.
16. Górnicka D, Kochańska J, Burduk A. Production Resources Utilization Improvement with the Use of Simulation Modelling. Advances in Intelligent Systems and Computing 2020; 1051: 41-50, https://doi.org/10.1007/978-3-030-30604-5_4.

17. Gunasekran A, Goyal SK, Martikainen T, Yli-Olli P. Production capacity planning and control in multi-stage manufacturing. Journal of the Operational Research Society 1998; 49(6): 625-634, https://doi.org/10.1038/sj.jors.2600580.
18. Hu Y, Guan Y, Han J, Wen J. Joint optimization of production planning and capacity adjustment for assembly system. Procedia CIRP 2017; 62: 193-198, https://doi.org/10.1016/j.procir.2016.06.029.
19. Jasiulewicz-Kaczmarek M, Gola A. Maintenance 4.0 Technologies for Sustainable Manufacturing - an Overview, IFAC PapersOnLine 2019; 52-10: 91-96, https://doi.org/10.1016/j.ifacol.2019.10.005.
20. Kłos S, Patalas-Maliszewska J. Using the simulation method for modelling a manufacturing system of predictive maintenance. Advances in Intelligent Systems and Computing 2020; 1001: 57-64, https://doi.org/10.1007/978-3-030-23946-6_7.
21. Kłosowski G, Rymarczyk T, Kania K, Świć A, Cieplak T. Maintenance of industrial reactors supported by deep learning driven ultrasound tomography. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 138-147, https://doi.org/10.17531/ein.2020.1.16.
22. Koren Y, Shpitalni M. Design of Reconfigurable manufacturing systems. Journal of Manufacturing Systems 2010; 29(4): 130-141, https://doi.org/10.1016/j.jmsy.2011.01.001.
23. Koren Y, Wang W, Gu X. Value creation through design for scalability of reconfigurable manufacturing systems. International Journal of Production Research 2016, 55(5): 1227-1242, https://doi.org/10.1080/00207543.2016.1145821.
24. Litwin P, Mądziel M, Stadnicka D. Simulations of Manufacturing Systems: Applications in Achieving the Intended Learning Outcomes. IFIP Advances in Information and Communication Technology 2019; 568: 615-623, https://doi.org/10.1007/978-3-030-28464-0_54.
25. Loska A, Paszkowski W. SmartMaintenance - the concept of supporting the exploitation decision-making process in the selected technical network system. Advances in Intelligent Systems and Computing 2018; 637: 63-47, https://doi.org/10.1007/978-3-319-64465-3_7.
26. Luss H. Operation research and capacity expansion problems: a survey. Operations Research 1982; 30(5), https://doi.org/10.1287/opre.30.5.907.
27. Maganha I, Silva C, Ferreira LMDF. Understanding reconfigurability of manfuacturing systems: An empirical analysis. Journal of Manufacturing Systems 2018; 48: 120-130, https://doi.org/10.1016/j.jmsy.2018.07.004.
28. Manne AS. Investments for capacity expansion, size, location, and time-phasing, The MIT Press, Cambridge, MA, 1967.
29. Młynarski S, Pilch R. Smolnik M, Szybka J, Wiązania G. A model of an adaptive strategy of preventive maintenance of complex technical objects. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 35-41, https://doi.org/10.17531/ein.2020.1.5.
30. Moghaddam SK, Houshmand M, Fatahi Valilai O. Configuration design in scalable reconfigurable manufacturing systems (RMS); a case of single-product flow line (SPFL). International Journal of Production Research 2018; 56(11): 3932-3954, https://doi.org/10.1080/00207543.2017.1412531.
31. Plinta D, Krajcovic M. Product system designing with the use of digital factory and augumented reality technologies. Advances in Intelligent Systems and Computing 2015; 350: 187-196, https://doi.org/10.1007/978-3-319-15796-2_19.
32. Putnik G, Sluga A, ElMraghy H, Teti R, Koren Y, Tolio T, Hon B. Scalability in manufacturing systems design and operation: State-of-the-art and future developments roadmap. CIRP Annals - Manufacturing Technology 2013; 62: 751-774, https://doi.org/10.1016/j.cirp.2013.05.002.
33. Singh A, Gupta S, Asjad M, Gupta P. Reconfigurable manufacturing systems: journey and the road ahead. International Journal of System Assurance Engineering and Management 2017; 8(2): 1849-1857, https://doi.org/10.1007/s13198-017-0610-z.
34. Sobaszek Ł, Gola A, Świć A. Time-based machine failure prediction in multi-machine manufacturing systems. Eksploatacja i Niezawodnosc-Maintenance and Reliability 2020; 22(1): 52-56, https://doi.org/10.17531/ein.2020.1.7.
35. Son S-Y, Olsen TL, Yip-Hoi D. An approach to scalability and line balancing for reconfigurable manufacturing systems. Integrated Manufacturing Systems 2001; 12(7): 500-511, https://doi.org/10.1108/09576060110407815.
36. Spicer P, Koren Y, Shpitalni M, Yip-Hoi D. Design principles for machining system configurations. CIRP Annals - Manufacturing Technology 2002; 51(1): 275-280, https://doi.org/10.1016/S0007-8506(07)61516-9.
37. Stevenson WJ. Operations Management. McGraw-Hill Education, New York, 2021.
38. Terkaj W, Tolio T, Valente A. Focused Flexibility in Production Systems [in:] Elmaraghy H.A. (ed.), Changeable and Reconfigurable Manufacturing Systems. Springer Series in Advanced Manufacturing. Springer, London, 2009, https://doi.org/10.1007/978-1-84882-067-8_3.
39. Vavrik V, Gregor M, Grznar P, Mozol S, Schickerle M, Durica L, Marschall M, Bielik T. Design of manufacturing lines using the reconfigurability principle. Mathematics 2020; 8(8): 1227, https://doi.org/10.3390/math8081227.
40. Wang W, Koren Y. Scalability planning for reconfigurable manufacturing systems. Journal of Manufacturing Systems 2012; 31: 83-91, https://doi.org/10.1016/j.jmsy.2011.11.001.
41. Zhang R. Research on Capacity Planning under Stochastic Production and Uncertain Demand. Systems Engineering - Theory & Practice 2007; 27(1): 51-59, https://doi.org/10.1016/S1874-8651(08)60006-X.



Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Duan R-X, He J-J, Feng T, Huang S-J, Chen L. Effective sensor placement based on a VIKOR method considering common cause failure in the presence of epistemic uncertainty. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 253–262, http://doi.org/10.17531/ein.2021.2.5.

Effective sensor placement based on a VIKOR method considering common cause failure in the presence of epistemic uncertainty

Indexed by: Web of Science Group

Rong-Xing Duan[a]*, Jie-Jun He[a], Tao Feng[a], Shu-Juan Huang[a], Li Chen[a]

[a]School of Information Engineering, Nanchang University, Nanchang, 330031, P. R. China

Highlights

- A VIKOR method is proposed to choose the possible sensor locations.
- A sensor model is presented by using a priority AND gate in sensor placement.
- CCF has an incredible influence on the reliability-based sensor placement method.

Abstract

Owing to expensive cost and restricted structure, limited sensors are allowed to install in modern systems to monitor the working state, which can improve their availability. Therefore, an effective sensor placement method is presented based on a VIKOR algorithm considering common cause failure (CCF) under epistemic uncertainty in this paper. Specifically, a dynamic fault tree (DFT) is developed to build a fault model to simulate dynamic fault behaviors and some reliability indices are calculated using a dynamic evidence network (DEN). Furthermore, a VIKOR method is proposed to choose the possible sensor locations based on these indices. Besides, a sensor model is introduced by using a priority AND gate (PAND) to describe the failure sequence between a sensor and a component. All placement schemes can be enumerated when the number of sensors is given, and the largest system reliability is the best alternative among the placement schemes. Finally, a case study shows that CCF has some influence on sensor placement and cannot be neglected in the reliability-based sensor placement.

Keywords

dynamic fault tree; epistemic uncertainty; sensor placement; reliability criterion; diagnostic sensor model

Acronyms and Abbreviations

CCF	Common cause failure
DFT	Dynamic fault tree
DEN	Dynamic evidence network
DIF	Diagnostic Importance Factor
BIM	Birnbaum Importance Measure
RAW	Risk Achievement Worth
PAND	Priority AND gate
DBN	Dynamic Bayesian network
FIM	Fisher information matrix
EFI	Effective independence method
MGL	Multi Greek Letter
MESH	Multiple error shock model
DTBN	Discrete-time Bayesian network
FTA	Fault tree analysis
BPA	Basic probability assignment

Notations

$\underline{P_i(x)}$	Lower bound of the failure probability of a component *i*
$\overline{P_i(x)}$	Upper bound of the failure probability of a component *i*
λ_I	Independent failure rate
λ_c	Common failure rate
β	Proportion of the probability of CCF in the total failure probability
P_{ind}	Probability of independent failure
P_{ccf}	Probability of CCF
P_{ij}	Proportion of the i^{th} alternative on the j^{th} attribute
h_j	Entropy value of the j^{th} attribute
ω_j	Weight value of the j^{th} attribute
C_j^+	Maximum range of each attribute
C_j^-	Minimum range of each attribute
c_{ij}	The j^{th} attribute value of the i^{th} component
c_j^+	Positive ideal solution
c_j^-	Negative ideal solution
S_i	Group benefit value
R_i	Individual regret degree
Q_i	Compromise value

(*) Corresponding author.
E-mail addresses: R-X. Duan - duanrongxing@ncu.edu.cn, J-J. He - 411014419018@email.ncu.edu.cn, T. Feng - 416114417131@email.ncu.edu.cn, S-J. Huang - 411014418170@email.ncu.edu.cn, L. Chen - 411016620088@email.ncu.edu.cn

1. Introduction

Driven by the support from modern technology, industrial production systems are seeing more synthesized and intelligent mechanical equipment. Predictably, the equipment is characterized by high risk, long cycle and expensive cost, which has more rigorous standards on diagnosis and maintenance. Therefore, it is particularly essential to avoid failures or locate the fault promptly when failures occur. Sensors are added to monitor the important components in the system, which not only provide early warning information to avoid major economic losses but also improve the efficiency of diagnosis when a fault occurs. The failure of the sensor to respond accurately matters much to the entire life of the sensor, which will escalate the difficulty of operation of the related equipment and make it delicate to satisfy specific environmental requirements. Under the assumption that the sensor will not fail, a sensor monitoring model constructed by static logic gates is given, and the sensor is added outside the structure of a fault tree [2]. Obviously, this model is no longer in step with the reality. The addition of sensors is bound to affect the reliability of the monitored system. To improve this, sensors are directly positioned on the monitored components in the concept of information fusion method [8], in effect diagnosing system fault using DFT analysis and DEN. However, the thorny problem of epistemic uncertainty remains unsolved and this approach has no access to consider that the addition of sensors will impact the system reliability. In reference [28], The sensor is taken as a component added in this system. A logic AND gate is adopted to describe the relationship between the component failure and sensor failure. When both failures occur, a failure will be output. However, this sensor monitoring model is not only easy to cause false alarms and increase the frequency of system maintenance unnecessarily, but also ignores missed alarms caused by the sequence of sensor failures and component failures. Hence, proposed by references [7, 11, 27], PAND gates are used to describe the time sequence between sensor failures and component failures. The Monte Carlo simulation and dynamic Bayesian network (DBN) are adopted to analyze DFT, which can effectively solve the above problems. Nevertheless, static fault tree is used to build the fault model and fails to describe the dynamic fault behaviors.

In the monitoring process of system status, the acquisition of system status dramatically depends on the effective sensor placement. The placement of sensors affects the monitoring capability of the sensor and the performance of the system. The location, type and quantity of sensors are major indices that determine the functionality, cost advantage and effectivity of sensor networks [28]. To assess the effectiveness of the sensor configurations, similarity of sensor locations and sensor distribution are usually taken into account [36]. The main goal of effective sensor placement is to select a set of sensor locations from a larger candidate set based on some available criteria. The Fisher information matrix (FIM) is used to give the solution of sensor placement for on-orbit modal identification and correlation of large space structures [15]. At the heart of FIM is to start from all possible monitoring positions, calculate the information matrix of each position and select the information matrix with the largest trace as the final position of the sensor. For this purpose, an optimal sensor placement is performed using the FIM [12]. On the other hand, an effective independence method (EFI) for optimal sensor placement is developed by using the FIM by Kammer [16]. Subsequently, the EFI method gains the growing popularity in the aspect of the best sensor placement [3, 5]. To achieve the goal of maximizing the effective information matrix determinant, a novel optimization of sensor placement is proposed using random EFI in reference [18]. The information matrix-based sensor placement method usually needs to decompose the eigenvalue of the matrix and calculate the inverse of the matrix. The calculation process is complicated and inefficient. Considering that the reduction of the modal assurance criterion has access to fewer iteration in sensor placement, a new multi-dimensional sensor placement criterion is presented by Yi [38] and a distributed wolf algorithm in the context of the paper is introduced to improve computational efficiency. Aiming at the defects of low modal energy and long calculation time of the modal matrix, a new modal shape matrix, established by He et al.[14], can overcome the above limitations. In reference [24], the locations of sensors are selected by minimizing information entropy, which is suited to assess the feasibility of sensor placement schemes in different forms. An optimization method based on information entropy, developed by Chow et al.[4], determines the sensor position of a typical power transmission tower with the updated structural model. Model-based optimization rules that consider diagnosable and cost constraints are another commonly used optimization method. Under certain condition of the known number of sensors, Duan [9] sets the objective function of the optimal sensor placement as the minimum expected diagnostic cost to resolve the sensor placement by the expected diagnostic cost, but ignoring sensor reliability. Xie et al.[35] presents an optimization strategy of the sensor placement, seeking the effective sensor placement by minimizing the average coherence while meeting budget constraints. Based on a hybrid model and data-driven method, a more effective and lower cost diagnosis and placement scheme in the system is presented by Zhang et al.[41]. It can quickly detect and locate the leakage area of the water-supply system. Steffelbauer et al.[33] incorporates different types and sources of uncertainty into the leak location of optimal sensor placement. For different numbers of sensors, the uncertainty of different intensities is considered. In addition, in order to depict the relationship between the number of sensors and the quality of leak location, a cost-benefit function is introduced using the different sensor placement results and GoF statistics. Generally, these methods are only suitable to specific systems. In fact, optimization algorithm is an issue that should be taken seriously during the process of optimizing sensor placement. Non-linear programming [31] is also widely used optimization method, but it is tempting to get a locally optimal solution. Targeting the above flaws cited, some optimization algorithms, such as genetic algorithms [37] and hybrid firefly algorithm with particle swarm optimization [25], are gaining the growing popularity in the domain of sensor placement. Arguably, the construction of a sensor model should be emphasized, a noteworthy problem in sensor placement. In reference [27], from the perspective of system fault diagnosis, a PAND gate is used to establish the sensor model and importance parameters of components are calculated to determine the potential sensor locations. Finally, the scheme with the least probability of system failure is the best sensor placement scheme. The above methods are essentially based on single-attribute decision-making, and the decision-making ability is not enough precise. For the placed object, the reliable and precise placement can be made by comprehensively considering multi-dimensional information. For this reason, in reference [28], a combination criterion based on the sensor failure risk and uncertainty of sensor information is developed to determine the effective placement of sensors, providing decision support for system health monitoring.

For the purpose of high reliability, some redundancy techniques are used in complex systems and make CCF exist when these systems break down. For the CCF problem, many scholars at home and abroad have established multiple CCF models, including the α-factor model [21], the β-factor model [17], the Multi Greek Letter (MGL) model [20] and the multiple error shock model (MESH) [19]. In reference [32], under the premise of considering CCF, a discrete-time Bayesian network (DTBN) is proposed to analyze the system reliability. Interval number theory is used for epistemic uncertainty and Matlab software is applied to calculate the reliability parameter. The β-factor model is built to handle the CCF problems, which converts static logic gates into DTBNs for analysis. Aiming at the epistemic uncertainty, a new sensor placement is proposed by using a DEN in reference [7], ignoring the CCF problem caused by simultaneous failure of blades and partitions in steam turbines due to high temperatures. In reference [44], an evidence network model is proposed to deal with the uncertainty of modal parameters and CCF. On this basis, the concept of multi-common cause failure and processing method is proposed [23].

According to the research of sensor placement mentioned above, most methods neglect CCF, epistemic uncertainty or dynamic fault behaviors. Additionally, a single indicator is used to choose the possible sensor locations, which will affect the effectiveness of sensor placement. This paper proposes a new effective sensor placement method to improve the effectiveness of sensor placement based upon the reliability criterion considering CCF problem and epistemic uncertainty shown as Fig. 1. A DFT is utilized to develop a fault model to simulate the dynamic fault behaviors. Besides, some reliability indices are calculated by mapping a DFT into a DEN, which can effectively handle CCF and solve the DFT with interval failure rate of components. Furthermore, a VIKOR-based method for determining the potential locations of the sensors is proposed based on multiple reliability parameters. Additionally, a sensor model is presented by using a priority AND gate (PAND) to describe the failure sequence between a sensor and a component. Finally, all placement schemes can be enumerated when the number of sensors is given, and the largest system reliability is the best alternative.

The remainder of this paper continues as follows. Section 2 focuses on the model construction of complex systems and solution for DFT considering CCF and epistemic uncertainty. An effective VIKOR method is developed to choose the possible sensor positions in section 3. Section 4 proposes a new sensor model to consider the failure sequence between components and sensors. The optimization of sensor placement is also proposed based on the optimal reliability criterion in Section 4. In Section 5, an ATP system is given to evaluate the effectiveness of the proposed method. Finally, some conclusions are made in Section 6.

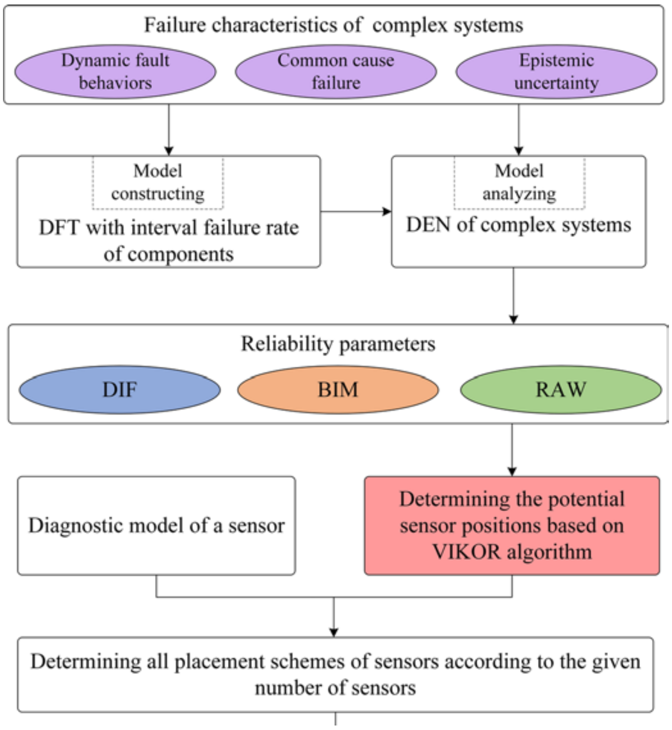


Fig. 1. Effective sensor placement method based on a VIKOR algorithm considering common cause failure in the presence of epistemic uncertainty

2. Reliability analysis based on DEN considering CCF

2.1. Construction of DFT Model

A fault tree [10] is a logical causal diagram representing the interactions between the components in a system when a failure occurs. In the fault tree, a series of specific logic gate symbols and transferring symbols are generally used to describe the causal relationship between various fault events and normal events in the system. Quantitative reliability and safety analysis are responsible for the growing acceptance of the fault tree analysis (FTA) [13]. The analysis is introduced to calculate the occurrence probability of the top event and recognize some important events in order to improve the system reliability. The traditionally static fault tree mostly includes some static logic gates. It is far from easy for the traditional static fault tree to describe the dynamic fault behaviors. In order to address this problem, the concept of DFT is developed by adding some dynamic logic gates based on the traditional fault tree approach. These dynamic logic gates generally include functional dependency gate, priority gate, sequential gate and spare gate. DFT can describe dynamic failure behaviors and are suited to evaluate the reliability of complex systems. In this paper, interval numbers are used to describe the failure rates of components based upon some datasheet over the period of product design.

2.2. Solution for DFT based on DEN under epistemic uncertainty

2.2.1. DEN

For two-state systems, all events only have two states: "occur" (F) and "not occur" (W). Accordingly, the knowledge framework of a component is $\Theta = \{F, W\}$ in evidence theory [6, 30], and all focal elements are defined as follows:

$$2^{\Theta} = \{\varnothing, \{F_i\}, \{W_i\}, \{F_i, W_i\}\} \tag{1}$$

where $\{F_i\}$ and $\{W_i\}$ respectively represent the fault state and normal state of a component or system, and $\{F_i, W_i\}$ represents the epistemic uncertainty.

Belief Function (Bel) represents the lower bound of the probability that the focus element exists, and Plausibility Function (Pl) represents the upper bound of the probability that the focus element exists. Accordingly, the basic probability assignment (BPA) of a component i is calculated as follows:

$$\begin{cases} m(\{W_i\}) = \mathrm{Bel}(\{W_i\}) \\ m(\{F_i\}) = 1 - \mathrm{Pl}(\{W_i\}) \\ m(\{F_i, W_i\}) = \mathrm{Pl}(\{W_i\}) - \mathrm{Bel}(\{F_i\}) \end{cases} \tag{2}$$

Evidence network, a widely used uncertainty reasoning method, has the advantages of D-S evidence theory and Bayesian network. It can more effectively solve the uncertainty problem of complex systems. DEN, an extension of initial evidence network in time, is a graphic structure and includes the original initial network and the time transfer network, where each time segment corresponds to a static evidence network. Each time segment is composed of a directed acyclic graph G_T=<V_T, E_T> and conditional probabilities, where V_T and E_T are represented as node sets and directed edge sets of time T respectively. Each time segment is connected by directed edges which are called transfer networks. In DEN, the state of the current time segment T depends only on the current state and the previous time segment T-ΔT, and has no relation with other states. The state of the current time segment T should meet the following requirements:

$$P(G_T \mid G_{T-\Delta T}, \ldots, G_{T_0}) = P(G_T \mid G_{T-\Delta T}) \tag{3}$$

However, the conditional belief distribution for the current focal element X with time k and the next focal element X with time k+1 should meet the following requirements:

$$m(X_{k+1} \mid X_k) = \begin{bmatrix} m(G_1^{X_{K+1}} \mid G_1^{X_K}) & \cdots & m(G_Q^{X_{K+1}} \mid G_1^{X_K}) \\ \vdots & \ddots & \vdots \\ m(G_1^{X_{K+1}} \mid G_Q^{X_K}) & \cdots & m(G_Q^{X_{K+1}} \mid G_Q^{X_K}) \end{bmatrix} \tag{4}$$

Table 1. The conditional mass distribution tables of node B(T+ΔT)

$B(T)$	$B(T+\Delta T)$		
	$\{W\}$	$\{F\}$	$\{W, F\}$
$\{W\}$	$m_B(W)$	$m_B(F)$	$m_B(W, F)$
$\{F\}$	0	1	0
$\{W, F\}$	0	$m_B(F)$	$1-m_B(F)$

2.2.2. Conversion of DFT into DEN

Static logic gates are majorly composed of AND gate, OR gate, and voting gate. The AND gate and PAND gate are applied to demonstrate the conversion of DFT into DEN in the following section. A logic AND gate outputs if any input event fails among the logical AND gate. A logic AND gate and the corresponding DEN are given in Fig. 2. The conditional probability table of node $B(T+\Delta T)$ in DEN is shown in Table 1 [22]. Formula (5) can be obtained from formula (2), showing the BPA of node B, and the conditional mass distribution formula of node $C(T+\Delta T)$ is given by formula (6).

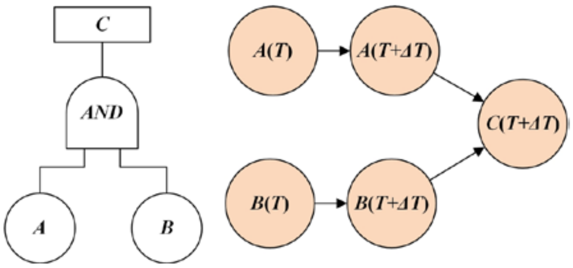


Fig. 2. A logic AND gate and the equivalent DEN

$$\begin{cases} m_B(W) = \exp(\bar{\lambda} \cdot \Delta T) \\ m_B(F) = 1 - \exp(\underline{\lambda} \cdot \Delta T) \\ m_B(F,W) = \exp(\underline{\lambda} \cdot \Delta T) - \exp(\bar{\lambda} \cdot \Delta T) \end{cases} \tag{5}$$

$$\begin{cases} P(C(T+\Delta T) = \{F\} \mid A(T+\Delta T) = \{F\}, B(T+\Delta T) = \{F\}) = 1 \\ P(C(T+\Delta T) = \{F\} \mid else) = 0 \\ P(C(T+\Delta T) = \{F,W\} \mid A(T+\Delta T) = \{W\}, B(T+\Delta T) = \{F,W\}) = 1 \\ P(C(T+\Delta T) = \{F,W\} \mid A(T+\Delta T) = \{F,W\}, B(T+\Delta T) = \{F\}) = 1 \\ P(C(T+\Delta T) = \{F,W\} \mid A(T+\Delta T) = \{F,W\}, B(T+\Delta T) = \{F,W\}) = 1 \\ P(C(T+\Delta T) = \{F,W\} \mid else) = 0 \end{cases} \tag{6}$$

The model of the PAND gate in the DEN is given in Fig. 3. The conditional probability table of node $A(T+\Delta T)$ is shown in Table 1. By using equations (7) and (8), the conditional probability formulas of the node $E(T+\Delta T)$ and $C(T+\Delta T)$ are obtained.

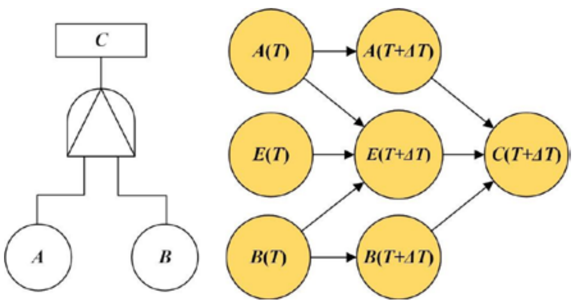


Fig. 3. A PAND gate and the equivalent DEN

$$\begin{cases} P(E(T+\Delta T) = \{F\} \mid E(T) = \{F\}) = 1 \\ P(E(T+\Delta T) = \{F\} \mid E(T) = \{W\}, A(T) = \{F\}) = 1 \\ P(E(T+\Delta T) = \{F\} \mid E(T) = \{F,W\}, A(T) = \{F\}) = 1 \\ P(E(T+\Delta T) = \{F\} \mid else) = 0 \\ P(E(T+\Delta T) = \{F,W\} \mid E(T) = \{F,W\}, A(T) = \{W\}) = 1 \\ P(E(T+\Delta T) = \{F,W\} \mid E(T) = \{W\}, A(T) = \{F,W\}) = 1 \\ P(E(T+\Delta T) = \{F,W\} \mid E(T) = \{F,W\}, A(T) = \{F,W\}) = 1 \\ P(E(T+\Delta T) = \{F,W\} \mid else) = 0 \end{cases} \tag{7}$$

$$\begin{cases} P(C(T+\Delta T) = \{F\} \mid E(T+\Delta T) = \{F\}, A(T+\Delta T) = \{F\}) = 1 \\ P(C(T+\Delta T) = \{F\} \mid else) = 0 \\ P(C(T+\Delta T) = \{F,W\} \mid E(T+\Delta T) = \{F,W\}, A(T+\Delta T) = \{W\}) = 1 \\ P(C(T+\Delta T) = \{F,W\} \mid E(T+\Delta T) = \{F,W\}, A(T+\Delta T) = \{F\}) = 1 \\ P(C(T+\Delta T) = \{F,W\} \mid else) = 0 \end{cases} \tag{8}$$

2.3. DEN model considering CCF

Redundant structure is usually used in complex systems to improve their performance. It is common that correlated failures often cause these systems to break down. If these correlated failures are ignored, it will lead to a big deviation in the reliability evaluation. CCF, one of the most common correlated failures, attracts more attention nowadays, and many researchers focus on this topic. CCF [43] is the simultaneous failure of two or more components due to some common causes. Explicit and implicit modeling methods are usually implemented to solve the CCF problem in reliability analysis [39]. The key to modeling a CCF system using DEN model is to make the component with CCF equivalent to an independent failure sub-component and a CCF sub-component, that is, the failure rate of CCF components in the system is divided into independent failure rate λ_I and CCF failure rate λ_c. The logical structure of the independent failure sub-component and the CCF sub-component is in series, and the common cause component failure occurs when any sub-component fails. Accordingly, in the DEN, the common cause event is regarded as the basic event of the system, that is to add a layer of independent failure sub-nodes and CCF sub-nodes on the basis of the root node, determine the edge probability of each sub-node, derive the conditional probability between each failure sub-node and components, and then construct the DEN model considering CCF. This paper adopts a β factor model to deal with CCF in the DEN. A network node without time change is added in the DEN, and its initial state is determined by the β factor value, as shown in Fig. 4.

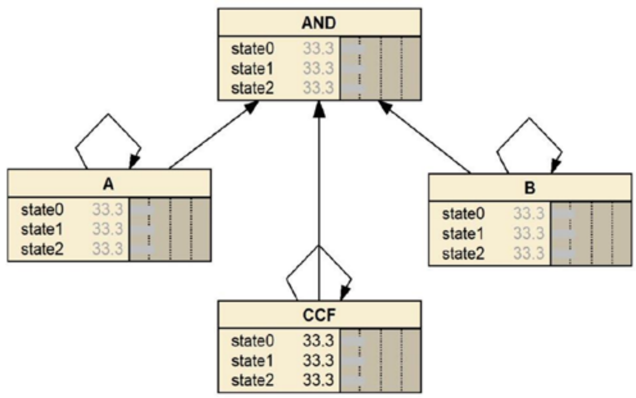


Fig. 4. An explicit modeling of AND gate considering CCF in the DEN

Generally, the parameter β can be defined as the proportion of the probability of CCF in the total failure probability. If a component obeys the exponential distribution, and the independent failure rate and the β-factor value are given, common failure rate can be calculated by the following equation.

$$\beta=\frac{1-e^{-\lambda_c t}}{1-e^{-\lambda_s t}}=\frac{1-e^{-\lambda_c t}}{(1-e^{-\lambda_I t})+(1-e^{-\lambda_c t})} \tag{9}$$

where λ_I is the independent failure rate of the component; λ_c is the CCF rate; λ_s is the whole failure rate of the component.

When the independent failure rate of the component is expressed by an interval number $[\underline{\lambda_I},\overline{\lambda_I}]$, the interval CCF rate $[\underline{\lambda_c},\overline{\lambda_c}]$ of components can be obtained according to the following formula:

$$\beta=\frac{1-\mathrm{e}^{-\underline{\lambda_c}\mathrm{t}}}{(1-\mathrm{e}^{-\underline{\lambda_c}\mathrm{t}})+(1-\mathrm{e}^{-\underline{\lambda_I}\mathrm{t}})} \tag{10}$$

$$\beta=\frac{1-\mathrm{e}^{-\overline{\lambda_c}\mathrm{t}}}{(1-\mathrm{e}^{-\overline{\lambda_c}\mathrm{t}})+(1-\mathrm{e}^{-\overline{\lambda_I}\mathrm{t}})} \tag{11}$$

The value of β usually range from 0 to 0.25. Actual components and the corresponding CCF influence should be considered to determine the specific value of β.

2.4. Calculating reliability results

Once the DFT model of a system is constructed, DFT is converted into the corresponding DEN based on the above approach. Some inference algorithms for DEN are applied to calculate some reliability indices. Three reliability parameters of DIF, BIM and RAW can be employed to quantify the influence of component on system reliability. However, each parameter has its unique characteristics. DIF [29] can describe the contribution of component failure to system failure. BIM [26] is defined as the influence of a failed component on the system and it has nothing to do with the reliability of the component, and only depends on the reliability of other components and the structure of the system. In general, RAW [40] is defined as the ratio of the risk metric value obtained when a component fails at the base case value of the risk metric. It is used to estimate the risk achievement of the system failure caused by a component failure and represents the significance of keeping a component at the current level of reliability.

3. Determining the possible sensor positions based on a VIKOR algorithm

This section proposes a method to determine the potential positions of sensors using VIKOR-based method under epistemic uncertainty [1]. The specific flow chart is shown in Fig. 5.

3.1. Constructing the decision matrix

The evaluation object is a component in the system in the process of selecting potential locations. Then, each component represents an evaluation scheme, which is shown by set $C = \{C_1, C_2, \ldots, C_m\}$. The reliability parameter of a component can be used as an evaluation attribute (evaluation indicator), which is represented by set $v = \{v_1, v_2, \ldots, v_n\}$. The weight vector of is $\omega=\{\omega_1, \omega_2, \ldots \omega_n\}$, where ω_j is the corresponding weight value of the evaluation attribute v_j. An original decision matrix composed of m evaluation schemes and n evaluation attributes can be expressed by the following formula:

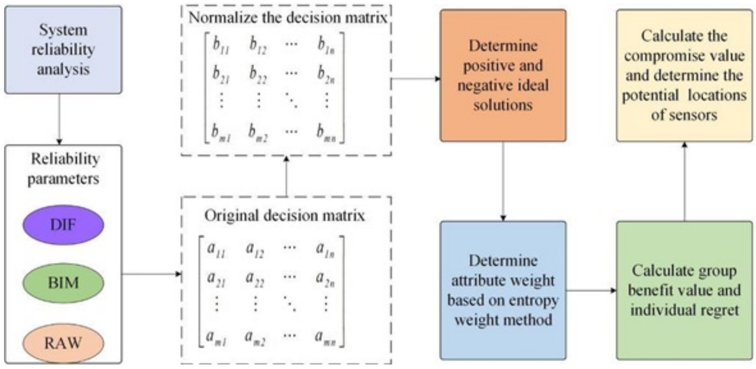


Fig. 5. A VIKOR-based method for determining the potential locations of sensors under epistemic uncertainty

$$C=\left(c_{ij}\right)_{m\times n}=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} \tag{12}$$

3.2. Calculating the weights of attributes using an entropy weight approach

Step 1. Standardize the decision matrix to tackle the homogenization of attributes' values. The negative and positive indexes can be calculated by the following equations:

$$c_{ij}{}'=\frac{c_{ij}-\min(c_j)}{\max(c_j)-\min(c_j)} \tag{13}$$

$$c_{ij}{}'=\frac{\max(c_j)-c_{ij}}{\max(c_j)-\min(c_j)} \tag{14}$$

where $\min(c_j)$ and $\max(c_j)$ are the minimum and maximum value of the j^{th} index respectively.

Step 2. Calculate P_{ij} using the following equation:

$$P_{ij}=\frac{c_{ij}{}'}{\sum_{i=1}^{m} c_{ij}{}'} \tag{15}$$

where P_{ij} is the proportion of the i^{th} alternative on the j^{th} attribute.

Step 3. Entropy values of attributes can be obtained as follows:

$$e_j=-K\sum_{i=1}^{m} p_{ij}\ln p_{ij} \tag{16}$$

where $K=1/\ln m$, ($K>0$,$0\le P_{ij}$).

Step 4. Weight values of attributes can be calculated by the following equation:

$$\omega_j=\frac{1-e_j}{\sum_{j=1}^{n} e_j} \tag{17}$$

Using the above four steps, the weight matrix $\omega=\{\omega_1, \omega_2, \dots \omega_n\}$ of attributes can be obtained, and ω satisfies the following formula $\sum_{j=1}^{n} \omega_j = 1$, $0 \le \omega_j \le 1$.

3.3. Determining the possible locations of sensors using a VIKOR method

The steps of determining the possible locations of sensors are given as follows based on the VIKOR algorithm.

Step 1. Construct the decision matrix $C=(c_{ij})_{m\times n}$, where $c_{ij} = [c_{ij}^-, c_{ij}^+]$ is the j^{th} attribute value of the i^{th} component in the system. The specific process is shown in formula (12).

Step 2. Determine the range of each attribute value:

$$\begin{cases} C_j^+ = \max\limits_{1\le i\le m} \{c_{ij}^+\} \\ C_j^- = \min\limits_{1\le i\le m} \{c_{ij}^-\} \end{cases} \tag{18}$$

Step 3. For the attributes described in interval numbers, the following two formulas can be used to calculate the positive ideal solution c_j^+ and negative ideal solution c_j^- of the attribute respectively:

$$c_j^+ = [\underline{c_j^+}, \overline{c_j^+}] = \begin{cases} \left[\max\limits_{1\le i\le m}\{c_{ij}^-\}, \max\limits_{1\le i\le m}\{c_{ij}^+\}\right], [c_{ij}^-, c_{ij}^+] \text{ is a benefit attribute} \\ \left[\min\limits_{1\le i\le m}\{c_{ij}^-\}, \min\limits_{1\le i\le m}\{c_{ij}^+\}\right], [c_{ij}^-, c_{ij}^+] \text{ is a cost attribute} \end{cases} \tag{19}$$

$$c_j^- = [\underline{c_j^-}, \overline{c_j^-}] = \begin{cases} \left[\min\limits_{1\le i\le m}\{c_{ij}^-\}, \min\limits_{1\le i\le m}\{c_{ij}^+\}\right], [c_{ij}^-, c_{ij}^+] \text{ is a cost attribute} \\ \left[\max\limits_{1\le i\le m}\{c_{ij}^-\}, \max\limits_{1\le i\le m}\{c_{ij}^+\}\right], [c_{ij}^-, c_{ij}^+] \text{ is a benefit attribute} \end{cases} \tag{20}$$

Step 4. Original decision matrix C can be normalized based on the Hamming distance, and the normalized decision matrix $B = (b_{ij})_{m\times n}$ is calculated by using the following equation:

$$b_{ij} = \frac{\left|\underline{c_j^+} - c_{ij}^-\right| + \left|\overline{c_j^+} - c_{ij}^+\right|}{2(C_j^+ - C_j^-)} \tag{21}$$

Step 5. Apply formula (17) to get the weight matrix $\omega=\{\omega_1, \omega_2, \dots \omega_j \dots \omega_n\}$, where $\omega_1 + \omega_2 + \cdots \omega_n = 1$, $\omega_j \in [0,1]$.

Step 6. Calculate the group benefit value S_i, the individual regret degree R_i and the compromise value Q_i:

$$S_i = \sum_{j=1}^{n} \omega_j \cdot b_{ij} \tag{22}$$

$$R_i = \max_{1\le j\le n} \{\omega_j \cdot b_{ij}\} \tag{23}$$

$$Q_i = v\frac{S_i - S^+}{S^+ - S^-} + (1-v)\frac{R_i - R^+}{R^+ - R^-} \tag{24}$$

where S^+ and S^- are the maximum and minimum values of group benefit S_i respectively; R^+ and R^- are the maximum and minimum values of individual regret R_i respectively. v is a constant. This paper assumes v=0.5, which means that maximizing group benefits is worthwhile minimizing group individual regret. The compromise value Q_i is sorted in ascending order. An equivalent number of system components or nodes with ranking among the top in Q_i are selected as the possible locations of sensors in light of the number of sensors.

4. Sensor placement method using reliability criterion

4.1. Sensor model

Some sensors are installed to monitor the operation state of some components in modern systems. When the value detected by a sensor is above the threshold, the sensor will give the alarm to the maintenance staff to repair or replace the component. Nevertheless, if a component fails after a sensor, and the monitored value is above the threshold, an alarm is not activated by this sensor until the component fails. In the following section, the temporal and logic relation will be described by using a new sensor model.

The output failure situation of the sensor monitoring model constructed in this paper has the following three situations.

(1) If the sensor does not fail before the monitored component fails, the sensor can monitor the state of component normally, and find the abnormal component in time to maintain or replace it. At this time, the entire model is considered normal.

(2) If the monitored component fails after the sensor fails, the sensor loses its function. At this time, the normality or failure of the entire model is determined by the working status of the detected component.

(3) The entire model is considered as a failure when the sensor and the monitored component fail at the same time.

A sensor is thought of as a component in a system in light of considering the reliability of this sensor. This paper uses the PAND gate to construct a sensor monitoring model based on the above discussion. This sequential failure can be captured by using a PAND gate, as shown in Fig. 6.

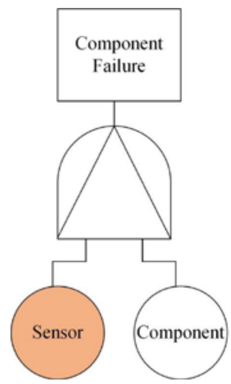


Fig. 6. A PAND gate to model the logic relation between a component and a sensor

4.2. Determining the optimal sensor placement scheme

Given the restrictions of structure and economic cost, only several sensors are allowed to be installed in some important locations. Let us suppose that the number of sensors is given. Usually, the number of locations detected is greater than the number of sensors. In this paper, there are M sensors installed in the system and N possible locations monitored by sensors (M < N), all possible placement schemes can be obtained using the following equation:

$$C(N,M) = \frac{N!}{M!(N-M)!} \tag{25}$$

For example, if there are only three allowed sensors to be placed in the system, X1, X2, X3 and X4 at the top of components, can be selected as the potential monitored positions of the sensors based on the described method for determining the potential position of the sensor. Assuming that there are four specific types of sensors S1, S2, S3 and S4 corresponding to four components, system will have the following four candidate placement schemes.

Scenario 1: Mount sensor S1 on component X1, mount sensor S2 on component X2, and mount sensor S3 on component X3.

Scenario 2: Mount sensor S1 on component X1, mount sensor S2 on component X2, and mount sensor S4 on component X4.

Scenario 3: Mount sensor S1 on component X1, mount sensor S3 on component X3, and mount sensor S4 on component X4.

Scenario 4: Mount sensor S2 on component X2, mount sensor S3 on component X3, and mount sensor S4 on component X4.

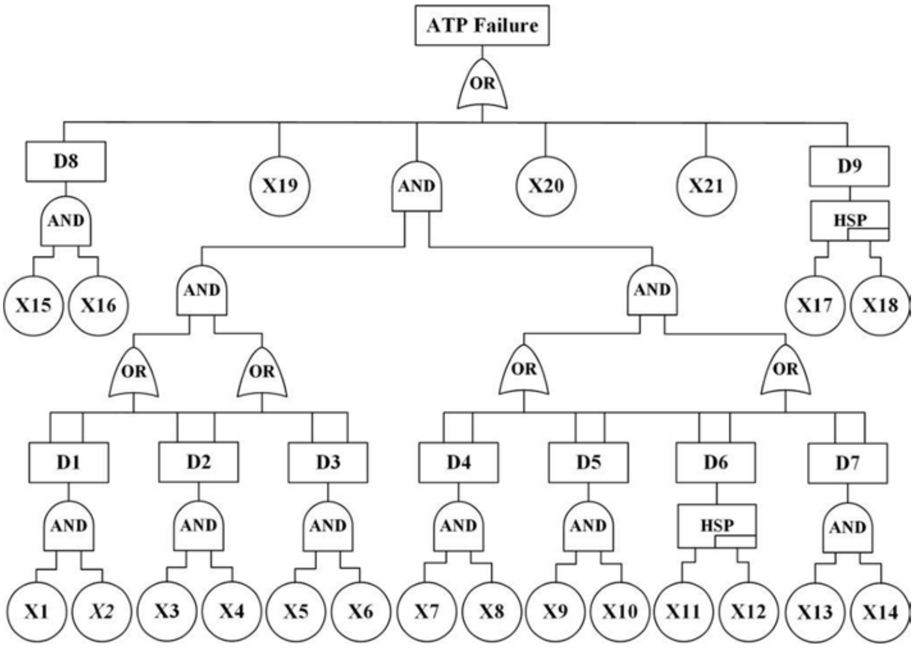


Fig. 7. A simplified fault tree model of ATP system

According to the proposed method, all possible placement scenarios can be obtained. A PAND gate, used to model the time dependences, is added to each scenario and the system reliability is calculated by the analysis of the updated DFT using the DEN based method. The best placement scheme is the scenario in which the system reliability is the largest.

5. A case study

The CTCS-3 ATP system [42] is a critical subsystem to guarantee the stable operation of trains and realize ultra-high-speed protection. Analyzing the reliability of the ATP system, finding out the key components or weak nodes of the system as potential installation locations of sensors, and optimizing the sensor placement scheme are of great significance to ensuring the safety of trains and reducing maintenance costs. The fault tree model of CTCS-3 ATP system is given in Fig. 7. Supposing that all components in the ATP system follow the exponential distribution and the failure rate of each component is expressed in the form of a definite value. In the presence of the epistemic uncertainty, the failure rate of the component is described in the form of interval numbers, as shown in Table 2.

To improve the reliability of ATP system, dual module redundant structure is used in the D1~D9 elements, and CCF exists in these modules. In this paper, a β-factor model is used to solve the problem of CCF. Under the condition that the independent failure rate λ_I of the component is given, and the interval failure rate $[\underline{\lambda_I},\overline{\lambda_I}]$ is obtained by formula $[\underline{\lambda_I},\overline{\lambda_I}]=[0.8\lambda_I,1.2\lambda_I]$. If β is known to be 10%, the CCF rate λ_c and interval CCF rate $[\underline{\lambda_c},\overline{\lambda_c}]$ can be obtained by formula (9), formula (10) and formula (11), as shown in Table 3.

The assumption is that the mission time T is 4000 hours and ΔT is 1000 hours. the DFT of ATP system can be converted into a DEN based on the approach mentioned above. In the two cases of considering CCF or not, the DEN is used to calculate DIF, BIM and RAW as the evaluation attributes. Two original decision matrices are given in Table 4 and Table 5. The entropy weight method determines the weight of each attribute as shown in Table 6. Table 7 shows the group benefit value S, individual regret R and the compromise value Q obtained by the VIKOR algorithm. Since interval numbers cannot be directly compared, then, the interval number ranking approach based on NSG possibility degree [34] is used to calculate the corresponding ranking values of BIM in Table 4 and Table 5, as shown in Table 8.

Table 2. Failure rates of all components in ATP system

Components	Failure rate λ_I /h	Interval failure rates $[\underline{\lambda_C},\overline{\lambda_C}]/h$
X1, X2	1.20e-5	[0.96e-5, 1.44e-5]
X3, X4	2.30e-6	[1.84e-6, 2.76e-6]
X5, X6	2.10e-5	[1.68e-5, 2.52e-5]
X7, X8	1.80e-5	[1.44e-5, 2.16e-5]
X9, X10	1.45e-8	[1.16e-8, 1.74e-8]
X11, X12	1.20e-5	[0.96e-5, 1.44e-5]
X13, X14	1.49e-5	[1.19e-5, 1.79e-5]
X15, X16	2.50e-9	[2.00e-9, 3.00e-9]
X17, X18	6.00e-6	[4.80e-6, 7.20e-6]
X19	2.00e-6	[1.60e-6, 3.20e-6]
X20	7.00e-8	[5.60e-8, 8.40e-8]
X21	5.00e-6	[4.00e-6, 6.00e-6]

Table 3. The interval CCF rates of all components in ATP system

Components	CCF failure rate λc /h	Interval CCF failure rate $[\underline{\lambda_I},\overline{\lambda_I}]/h$
X1, X2	1.20e-5	[1.02e-6,1.50e-6]
X3, X4	2.30e-6	[2.03e-7,3.03e-7]
X5, X6	2.10e-5	[1.73e-6,2.51e-6]
X7, X8	1.80e-5	[1.50e-6,2.18e-6]
X9, X10	1.45e-8	[1.29e-9,1.93e-9]
X11, X12	1.20e-5	[1.02e-6,1.50e-6]
X13, X14	1.49e-5	[1.25e-6,1.84e-6]
X15, X16	2.50e-9	[2.22e-10,3.33e-10]
X17, X18	6.00e-6	[5.22e-7,7.75e-7]

Assuming that only two sensors are allowed to be placed in the system, three nodes are designated as the potential sensor positions by the formula (24). Regardless of whether the CCF is considered, it

Table 4. *The original decision matrix ignoring CCF*

Nodes	DIF	BIM	RAW
D1	[0.0004, 0.0009]	[0, 0.0123]	[0.6714, 1.9919]
D2	[0, 0]	[0, 0.0123]	[0.6714, 1.9919]
D3	[0.0879, 0.0891]	[0.9814, 0.9887]	[47.6190, 80.6451]
D4	[0.0008, 0.0018]	[0, 0.0094]	[0.6095, 1.7581]
D5	[0, 0]	[0, 0.0094]	[0.6095, 1.7581]
D6	[0.0004, 0.0008]	[0, 0.0094]	[0.6095, 1.7581]
D7	[0.0006, 0.0013]	[0, 0.0094]	[0.6095, 1.7581]
D8	[0, 0]	[0.9790, 0.9876]	[47.6190, 80.6451]
D9	[0.0074, 0.0075]	[0.9792, 0.9877]	[47.6190, 80.6451]
X19	[0.2573, 0.2696]	[0.9853, 0.9907]	[47.6190, 80.6451]
X20	[0.0090, 0.0091]	[0.9792, 0.9877]	[47.6190, 80.6451]
X21	[0.6416, 0.6431]	[0.9909, 0.9955]	[47.6190, 80.6451]

Table 5. *The original decision matrix considering CCF*

Nodes	DIF	BIM	RAW
D1	[0.0037, 0.0051]	[0, 0.0248]	[0.9453, 2.4970]
D2	[0.0006, 0.0009]	[0, 0.0248]	[0.9453, 2.4970]
D3	[0.2691, 0.2711]	[0.9798, 0.9876]	[36.4964, 59.8802]
D4	[0.0044, 0.0067]	[0, 0.0148]	[0.7153, 1.9581]
D5	[0, 0]	[0, 0.0148]	[0.7153, 1.9581]
D6	[0.0028, 0.0042]	[0, 0.0148]	[0.7153, 1.9581]
D7	[0.0035, 0.0054]	[0, 0.0148]	[0.7153, 1.9581]
D8	[0, 0]	[0.9726, 0.9831]	[0.7153, 1.9581]
D9	[0.0672, 0.0678]	[0.9743, 0.9842]	[36.4964, 59.8802]
X19	[0.1893, 0.1918]	[0.9788, 0.9863]	[36.4964, 59.8802]
X20	[0.0066, 0.0067]	[0.9727, 0.9832]	[36.4964, 59.8802]
X21	[0.4720, 0.4741]	[0.9843, 0.9910]	[36.4964, 59.8802]

Table 6. *The weight table of attributes*

Weight	Ignoring CCF	Considering CCF
ω_{DIF}	0.1931	0.1899
ω_{BIM}	0.1931	0.1893
ω_{RAW}	0.6138	0.6208

Table 7. *S, R and Q values of each node*

Nodes	Ignoring CCF			Considering CCF		
	S	R	Q	S	R	Q
D1	0.8657	0.4817	0.9979	0.8616	0.4876	0.9920
D2	0.8659	0.4817	0.9980	0.8630	0.4876	0.9929
D3	0.1679	0.1663	0.2690	0.0821	0.0813	0.1299
D4	0.8670	0.4828	0.9998	0.8661	0.4916	0.9987
D5	0.8674	0.4828	1.0000	0.8683	0.4916	1.0000
D6	0.8672	0.4828	0.9999	0.8669	0.4916	0.9992
D7	0.8671	0.4828	0.9998	0.8665	0.4916	0.9990
D8	0.1948	0.1929	0.3120	0.6830	0.4916	0.8933
D9	0.1925	0.1906	0.3084	0.1641	0.1625	0.2597
X19	0.1148	0.1138	0.1840	0.1141	0.1132	0.1808
X20	0.1920	0.1901	0.3076	0.1887	0.1868	0.2987
X21	0	0	0	0	0	0

Table 8. *The sort value corresponding to BIM*

Nodes	Ignoring CCF	Considering CCF	Nodes	Ignoring CCF	Considering CCF
D1	0.0596	0.0605	D7	0.0580	0.0576
D2	0.0596	0.0605	D8	0.1043	0.1040
D3	0.1064	0.1106	D9	0.1044	0.1054
D4	0.0580	0.0576	X19	0.1108	0.1092
D5	0.0580	0.0576	X20	0.1044	0.1041
D6	0.0580	0.0576	X21	0.1184	0.1156

Table 9. *Failure rate of sensors*

Sensors	Failure rate λ_I/h	Interval failure rate $[\underline{\lambda_I}, \overline{\lambda_I}]/h$
S1	4.05e-7	[3.24e-7, 4.86e-7]
S2	9.30e-7	[7.44e-7, 11.16e-7]
S3	4.20e-6	[3.36e-6, 5.04e-6]

is painfully obvious that the compromise value Q of nodes D3, X19 and X21 is smaller in Table 7; The BIM of nodes D3, X19 and X21 correspond to larger ranking values are obtained in Table 8. Therefore, under the above conditions, these nodes are chosen as the possible positions of sensors in the ATP system. Suppose that sensors S1, S2 and S3 are specific types of sensors that monitor nodes X19, X21 and D3, respectively. The sensor monitoring model composed of PAND gates introduced in this paper is added to the system fault tree model, then all sensor placement schemes of the system are as follows.

Scheme 1: Install sensor S1 on node X19 and install sensor S2 on node X21.

Scheme 2: Install sensor S1 on node X19 and install sensor S3 on node D3.

Scheme 3: Install sensor S2 on node X21 and install sensor S3 on node D3.

Table 10. *The reliability of ATP system and the corresponding ranking value when failure rate is an interval value*

Schemes	System reliability ignoring CCF	Ranking value	System reliability considering CCF	Ranking value
1	[0.9972, 0.9987]	0.5000	[0.9905, 0.9942]	0.3864
2	[0.9877, 0.9958]	0.2043	[0.9860, 0.9907]	0.1706
3	[0.9932, 0.9966]	0.2957	[0.9916, 0.9955]	0.4430

The sensor, as a high-reliability component, is generally dozens of times lower than the failure rate of the monitored component. Therefore, it can be reasonably assumed that the sensor failure rate is given in Table 9. For the interval failure rate of node (component), the fault tree model of ATP

system can be mapped into the DEN to calculate the system reliability under various scenarios, or, the normal probability of the system at the end of the system task time. Table 10 gives the system reliability and its corresponding ranking values under various placement schemes when failure rate of the node (component) is interval number. It can conclude that the optimal sensor placement scheme in the ATP system ignoring CCF is scenario 1 and the optimal placement scheme considering CCF is scenario 3 according to Table 10. Considering whether CCF or not, the optimal placement scheme is different.

Hence, conclusions can be made that CCF generates an incredibly important impact on sensor placement using reliability criterion and cannot be neglected in sensor placement analysis.

6. Conclusion

This paper proposes an effective sensor placement method based on the reliability criterion in the presence of epistemic uncertainty. It is designed to tackle two important challenges emerging in complex systems, for example, CCF in components and dynamic fault behaviors. Aiming at the problem of CCF, the β-factor model is adopted to address the CCF failure rate and independent failure rate of components. For the issue of dynamic fault behaviors, a DFT is used to construct a fault model and the DFT is mapped into a DEN to compute several reliability indices used as evaluation attributes to build a decision matrix. Additionally, the potential locations of sensors are obtained using an efficient VIKOR algorithm and a diagnostic sensor model is constructed based on a PAND gate to capture the sequence between sensor failures and the monitored component failures. Furthermore, the best sensor placement scheme is obtained based on the system reliability among the placement schemes. Finally, an actual ATP system is given to evaluate the effectiveness of the proposed method. Some conclusions are made that CCF generates an incredibly important impact on sensor placement using reliability criterion and cannot be neglected in sensor placement analysis. The proposed method makes full use of the advantages of DFT for modeling, DEN for solving the problem of epistemic uncertainty and a VIKOR algorithm for decision making, which particularly is appropriate for effective sensor placement in complex engineering systems.

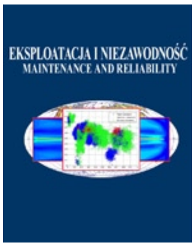
Acknowledgement

The research is partly supported by National Natural Science Foundation of China (71961017) and Graduate Innovation Foundation of Jiangxi Province (YC2019-S071).

References

1. Acuña-Soto CM, Liern V, Pérez-Gladish B. A VIKOR-based approach for the ranking of mathematical instructional videos. Management Decision 2019; 57(2): 501-522, http://doi.org/10.1108/md-03-2018-0242.
2. Assaf T, Dugan JB. Diagnosis based on reliability analysis using monitors and sensors. Reliability Engineering & System Safety 2008; 93(4): 509-521, http://doi.org/10.1016/j.ress.2006.10.024.
3. Błachowski B, Świercz A, Ostrowski M, Tauzowski P, Olaszek P, Jankowski Ł. Convex relaxation for efficient sensor layout optimization in large-scale structures subjected to moving loads. Computer-Aided Civil and Infrastructure Engineering 2020; 35(10): 1085-1100, http://doi.org/10.1111/mice.12553.
4. Chow HM, Lam HF, Yin T, Au SK. Optimal sensor configuration of a typical transmission tower for the purpose of structural model updating. Structural Control and Health Monitoring 2011; 18(3): 305-320, http://doi.org/10.1002/stc.372.
5. Dai G B, Ji G Y. Structure of Large Ground Modal Test Node Selecting and Optimizing Method Research. Journal of Vibration, Measurement & Diagnosis 2017; 37(05): 984-989, http://doi.org/10.16450/j.cnki.issn.1004-6801.2017.05.020.
6. Dempster A P. Upper and lower probabilities induced by a multivalued mapping. Classic works of the Dempster-Shafer theory of belief functions 2008; 57-72, http://doi.org/10.1007/978-3-540-44792-4_3.
7. Duan R, Lin Y, Feng T. Optimal Sensor Placement Based on System Reliability Criterion Under Epistemic Uncertainty. IEEE Access 2018; 6: 57061-57072, http://doi.org/10.1109/access.2018.2873420.
8. Duan R, Lin Y, Zeng Y. Fault diagnosis for complex systems based on reliability analysis and sensors data considering epistemic uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 558-566, http://doi.org/10.17531/ein.2018.4.7.
9. Duan R, Ou D, Dong D, Zhou H. Optimal sensor placement for fault diagnosis based on diagnosis cost specifications. Journal of Computational Information Systems 2011; 7(9): 3253-3260,
10. Dugan J B, Bavuso S J, Boyd M A. Dynamic fault-tree models for fault-tolerant computer systems. IEEE Transactions on reliability 1992; 42(3): 363-377, http://doi.org/10.1109/24.159800.
11. Feng T, Duan R, Lin Y, Zeng Y. Optimal sensor placement based on multiattribute decision-making considering the common cause failure. Proceedings of the Institution of Mechanical Engineers, Part C Journal of Mechanical Engineering Science 2019; 233(19/20): 7170-7182, http://doi.org/10.1177/0954406219864982.
12. Gomes GF, Pereira JVP. Sensor placement optimization and damage identification in a fuselage structure using inverse modal problem and firefly algorithm. Evolutionary Intelligence 2020, http://doi.org/10.1007/s12065-020-00372-1.
13. Han X, Yin X, Zhang Z. Application of fault tree analysis method in reliability analysis of substation communication system. Power System Technology 2004.
14. He C, Xing J, Li J, Yang Q, Wang R, Zhang X. A New Optimal Sensor Placement Strategy Based on Modified Modal Assurance Criterion and Improved Adaptive Genetic Algorithm for Structural Health Monitoring. Mathematical Problems in Engineering 2015; 2015: 1-10, http://doi.org/10.1155/2015/626342.
15. Kammer DC. Sensor placement for on-orbit modal identification and correlation of large space structures. Journal of Guidance, Control, and Dynamics 1991; 14(2): 251-259, http://doi.org/ 0.2514/3.20635.
16. Kammer DC, Tinker ML. Optimal placement of triaxial accelerometers for modal vibration tests. Mechanical Systems and Signal Processing 2004; 18(1): 29-41, doi:10.1016/s0888-3270(03)00017-7.
17. Kančev D, Čepin M. A new method for explicit modelling of single failure event within different common cause failure groups. Reliability Engineering & System Safety 2012; 103: 84-93, http://doi.org/10.1016/j.ress.2012.03.009.
18. Kim T, Youn BD, Oh H. Development of a stochastic effective independence (SEFI) method for optimal sensor placement under uncertainty. Mechanical Systems and Signal Processing 2018; 111: 615-627, http://doi.org/10.1016/j.ymssp.2018.04.010.
19. Li Z Q, Xu T X, An J, Fu L Y, Gu J Y. Common cause failure modeling for redundant system based on dynamic Bayesian network. Chinese Journal of Scientific Instrument 2018; 39(3): 190-198, http://doi.org/10.19650/j.cnki.cjsi.J1702575.
20. Li Z Q, Xu T X, Gu J Y, Liu Y D. Review on Research on Dependent Failure Analysis of Complex Systems. Failure Analysis and Prevention 2017; 12(2): 130-136.
21. Mi J, Li Y-F, Beer M, Broggi M, Cheng Y. Importance measure of probabilistic common cause failures under system hybrid uncertainty based on bayesian network. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22(1): 112–120, http://doi.org/10.17531/ein.2020.1.13.

22. Mi J, Li Y-F, Peng W, Huang H-Z. Reliability Analysis of Complex Multi-state System with Common Cause Failure Based on DS Evidence Theory and Bayesian Network. Recent Advances in Multi-state Systems Reliability. 2018: 19-38, http://doi.org/10.1007/978-3-319-63423-4_2.
23. Mi J, Li Y-F, Peng W, Huang H-Z. Reliability analysis of complex multi-state system with common cause failure based on evidential networks. Reliability Engineering & System Safety 2018; 174: 71-81, http://doi.org/10.1016/j.ress.2018.02.021.
24. Papadimitriou C, Lombaert G. The effect of prediction error correlation on optimal sensor placement in structural dynamics. Mechanical Systems and Signal Processing 2012; 28: 105-127, http://doi.org/10.1016/j.ymssp.2011.05.019.
25. Pitchaimanickam B, Murugaboopathi G. A hybrid firefly algorithm with particle swarm optimization for energy efficient optimal cluster head selection in wireless sensor networks. Neural Computing and Applications 2020; 32(12): 7709-7723,
26. Salazar JC, Weber P, Nejjari F, Sarrate R, Theilliol D. System reliability aware Model Predictive Control framework. Reliability Engineering & System Safety 2017; 167: 663-672, http://doi.org/10.1016/j.ress.2017.04.012.
27. Salehpour-Oskouei F, Pourgol-Mohammad M. Fault Diagnosis Improvement Using Dynamic Fault Model in Optimal Sensor Placement: A Case Study of Steam Turbine. Quality and Reliability Engineering International 2017; 33(3): 531-541, http://doi.org/10.1002/qre.2031.
28. Salehpour-Oskouei F, Pourgol-Mohammad M. Sensor placement determination in system health monitoring process based on dual information risk and uncertainty criteria. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2017; 232(1): 65-81, http://doi.org/10.1177/1748006x17742766.
29. Sallak M, Schon W, Aguirre F. Extended Component Importance Measures Considering Aleatory and Epistemic Uncertainties. IEEE Transactions on Reliability 2013; 62(1): 49-65, http://doi.org/10.1109/tr.2013.2240888.
30. Shafer G. A mathematical theory of evidence. Princeton University Press, 1976.
31. Shan J, Shi W, Lu X. Model-Reference Health Monitoring of Hysteretic Building Structure Using Acceleration Measurement with Test Validation. Computer-Aided Civil and Infrastructure Engineering 2016; 31(6): 449-464, http://doi.org/10.1111/mice.12172.
32. Song Y, Mi J, Cheng Y, Bai L, Wang X. Application of discrete-time Bayesian network on reliability analysis of uncertain system with common cause failure. Quality and Reliability Engineering International 2018; 35(4): 1025-1045, http://doi.org/10.1002/qre.2443.
33. Steffelbauer DB, Fuchs-Hanusch D. Efficient Sensor Placement for Leak Localization Considering Uncertainties. Water Resources Management 2016; 30(14): 5517-5533, http://doi.org/10.1007/s11269-016-1504-6.
34. Sun H, Yao W. Comments on methods for ranking interval numbers. Journal of Systems Engineering 2010; 25(3): 304-312.
35. Xie X, Zhou Q, Hou D, Zhang H. Compressed sensing based optimal sensor placement for leak localization in water distribution networks. Journal of Hydroinformatics 2018; 20(6): 1286-1295, http://doi.org/10.2166/hydro.2017.145.
36. Yang C, Ma R, Ma R. Optimal Sensor Placement for Modal Identification in Multirotary-Joint Solar Power Satellite. IEEE Sensors Journal 2020; 20(13): 7337-7346, http://doi.org/10.1109/jsen.2020.2978081.
37. Yang C, Zhang X, Huang X, Cheng Z, Zhang X, Hou X. Optimal sensor placement for deployable antenna module health monitoring in SSPS using genetic algorithm. Acta Astronautica 2017; 140: 213-224, http://doi.org/10.1016/j.actaastro.2017.08.025.
38. Yi T-H, Li H-N, Wang C-W. Multiaxial sensor placement optimization in structural health monitoring using distributed wolf algorithm. Structural Control and Health Monitoring 2016; 23(4): 719-734, http://doi.org/10.1002/stc.1806.
39. Yu G, Du Y, Yan L, Ren F. Stress-strength interference-based importance for series systems considering common cause failure. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(2): 241-252, http://doi.org/10.17531/ein.2020.2.7.
40. Zarei E, Azadeh A, Khakzad N, Aliabadi MM, Mohammadfam I. Dynamic safety assessment of natural gas stations using Bayesian network. Journal of Hazardous Materials 2017; 321: 830-840, http://doi.org/10.1016/j.jhazmat.2016.09.074.
41. Zhang Q, Zheng Y, Zhao M, Qi J. Leakage Zone Identification in Large-Scale Water Distribution Systems Using Multiclass Support Vector Machines. Journal of Water Resources Planning & Management 2016; 142(11): 04016042, http://doi.org/10.1061/(ASCE)WR.1943-5452.0000661.
42. Zhang Y, Yang J. Reliability Analysis on ATP System of CTCS-3 Based on Dynamic Bayesian Network. Journal of the China Railway Society 2017; 39(7): 79-86, http://doi.org/10.3969/j.issn.1001-8360.2017.07.012.
43. Zubair M, Amjad QMN. Calculation and updating of Common Cause Failure unavailability by using alpha factor model. Annals of Nuclear Energy 2016; 90: 106-114, http://doi.org/10.1016/j.anucene.2015.12.004.
44. Zuo L, Xiahou T, Liu Y. Evidential network-based failure analysis for systems suffering common cause failure and model parameter uncertainty. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 2018; 233(6): 2225-2235, http://doi.org/10.1177/0954406218781407.



Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Rosienkiewicz M. Artificial intelligence-based hybrid forecasting models for manufacturing systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 263–277, http://doi.org/10.17531/ein.2021.2.6.

Artificial intelligence-based hybrid forecasting models for manufacturing systems

Indexed by: Web of Science Group

Maria Rosienkiewicz[a]

[a]Wroclaw University of Science and Technology, Faculty of Mechanical Engineering, Centre for Advanced Manufacturing Technologies, ul. Łukasiewicza 5, 50-371 Wrocław, Poland

Highlights

- In the paper 4 new hybrid forecasting artificial intelligence-based models are proposed.
- A problem of defining explanatory variables when access to data is limited is addressed.
- Study fills in literature gap of hybrid forecasting application in manufacturing systems.
- The presented case studies cover production planning, maintenance and quality control.
- Algorithm for forecasting accuracy assessment and optimal method selection is presented.

Abstract

The paper addresses the problem of forecasting in manufacturing systems. The main aim of the research is to propose new hybrid forecasting models combining artificial intelligence-based methods with traditional techniques based on time series – namely: Hybrid econometric model, Hybrid artificial neural network model, Hybrid support vector machine model and Hybrid extreme learning machine model. The study focuses on solving the problem of limited access to independent variables. Empirical verification of the proposed models is built upon real data from the three manufacturing system areas – production planning, maintenance and quality control. Moreover, in the paper, an algorithm for the forecasting accuracy assessment and optimal method selection for industrial companies is introduced. It can serve not only as an efficient and costless tool for advanced manufacturing companies willing to select the right forecasting method for their particular needs but also as an approach supporting the initial steps of transformation towards smart factory and Industry 4.0 implementation.

Keywords

artificial neural network, support vector machine, extreme learning machine, hybrid forecasting, production planning, maintenance, quality control.

1. Introduction

In many enterprises the issue of making accurate business decisions often depends on the quality of demand forecasts for manufactured products. “Demand forecasting is crucial for decision making and operations in organisations” [38]. In the era of globalization, market uncertainty, and growing supply chain complexity the need for integrated and efficient planning increases [55]. Predicting future demand values provides the basis not only for proper production planning, but also for preparing precise material, financial and employee demand schedules. The proper resources management is a challenging task in every manufacturing company [25]. The accurate demand forecasting for the manufactured products allows reducing inventory and improving order indicators, whereas “inaccurate forecasts can be costly for company operations, in terms of stock-outs and lost sales, or over-stocking, while not meeting service level targets” [39]. Forecasting is widely used not only in production planning, but also in maintenance – it enables the companies to predict failures or demand for spare parts (examples of forecasting applications in maintenance-related problems include for instance time-based machine failure prediction in multi-machine manufacturing systems [75] or lifetime prediction of bearings or bearing-based systems [4]). In manufacturing processes the quality of the end product is in general defined by multiple critical outputs or responses and therefore, the efficient forecasting of quality is both critical and challenging for practitioners [72]. In fact in every single area of activity of a manufacturing company, for which it is possible to collect the appropriate dataset and it is necessary to make effective decisions regarding future operations, accurate prediction techniques should be implemented.

In [29] Hall discusses a number of cases presenting how improvement of forecasting influences profitability of companies – for example Hyundai Motors has reduced delivery time by 20% and increased inventory turns from 3 to 3.4, whereas Reynolds Aluminum has reduced forecasting errors by 2%, which in turn caused a reduction of 1 million pounds in inventory. Moreover, Unilever has reduced forecasting errors from 40% to 25%, which has brought multi-million dollar savings. SCI Systems on the other hand has reduced on-hand inventory by 15%, which resulted in annual savings of 180 million dollars. It is also worth mentioning that Virgin Atlantic Cargo – being one of the largest air freight operators in the world – has identified forecasting accuracy as of strategic importance to its operational efficiency, due to the reason that efficient predictions ensure to have the right resources available at the right place and time [37].

Another important area of forecasting implementation in manufacturing companies is spare parts management. According to Suomala et al. the impact of the spare parts business is significant in terms of

E-mail addresses: M. Rosienkiewicz - maria.rosienkiewicz@pwr.edu.pl

a company's profit [76]. "Strategically aligned and efficiently implemented spare parts logistics can differentiate a business from its competitors, lower costs, increase revenues, and thus help firms generate greater value for customers and ultimately increase profits"[80]. In consequence it can be stated that improvement of processes related to spare parts management is a matter of great importance for many industrial companies. Especially challenging is the issue of lumpy and intermittent demand forecasting – a sort of demand typical for spare parts, which can be observed in aerospace, automotive, mining and railway industry as well as in advanced manufacturing or electronics. Forecasting the failure rate of machines based on data obtained from the monitoring systems is an extremely important solution for maintenance departments, which goal is to minimize the number of failures. When weighing these considerations against industrial implementations results, it can be noticed that predictive maintenance reduces the time needed for planned machinery maintenance by 20-50%, equipment availability can be increased by 10-20%, whereas overall maintenance costs can be reduced by 5-10% [17].

In the era of Industry 4.0, which can be defined as "an integration of intelligent machines, systems and the introduction of changes in production processes aimed at increasing production efficiency and introducing the possibility of flexible product changes"[70], the issue of accurate forecasting becomes especially important. The impact of new technologies like Big Data, Industrial Internet of Things and Cloud Computing, which are considered to be the pillars of Industry 4.0, changes the way how manufacturing companies operate. Especially "the field of big data time series has dramatically evolved in the last years"[56]. Cloud Computing "has become a new type of Internet service because of its high scalability, flexibility, and cost-efficiency"[9]. What is more, "from an ICT point-of-view, during the last decay, "Data, Information and Knowledge" (DIK) became a central capital with a critical value. ICTs introduced huge changes in Knowledge Management (KM) and AI applications"[8]. Flexible Manufacturing Systems - representing an opportunity for shifting from fixed to customized production - when associated with computational technology they lead to industry of the future [47]. According to a report on Industry 4.0 by PricewaterhouseCoopers huge data volumes generated by control systems, which are currently used mainly to monitor the state of technological processes, in the future will enable predicting their behaviour and product quality parameters, as well as global production control. Therefore it is expected that manufacturing management processes will be subjected to major changes [88]. In consequence, currently, a number of manufacturing companies are facing the challenge of transforming into so-called smart factory or factory of the future. In Poland, as most companies are currently at the stage of the third industrial revolution, the process of implementing Industry 4.0 technologies is still ahead. As the increasing amount of collected data requires effective analytical tools [41], there still is a need to develop new ways and models enhancing this transformation process. As Frank at al. underline "the effective implementation of Industry 4.0 technologies is still a subject of research"[22].

A properly constructed database is a key aspect of effective forecasting. Collecting data of adequate quality is a prerequisite for building accurate models. Currently, due to advanced information and manufacturing technologies, companies have the opportunity to gather a huge amount of data that characterize the work of machines, their technical condition and production processes. These data can be collected, for instance through sensors that are installed in particular machines and devices. In consequence, production management based on digital data allows adding value in the areas of production and logistics mainly due to very precise forecasting of demand and making manufacturing process more flexible, reducing failures by implementing predictive models in maintenance as well as elimination of root causes of defects through intelligent quality assurance processes.

As an access to data is getting easier in manufacturing systems, the companies are willing to possess the knowledge hidden in the data and develop efficient predictions. On the other hand, obtaining appropriate data to build accurate forecasting models is still rather challenging – especially, when data characterizing explanatory (independent) variables are desirable. Often companies interested in effective forecasting face the problem of lack of available, reliable, complete and comparable statistical data. Therefore there still is a need to develop approaches allowing the companies to create a set of potential explanatory variables when access to data is limited and to develop new ways enhancing companies in transformation to Industry 4.0. To answer this need the Author proposes new hybrid forecasting models dedicated to manufacturing systems.

2. Literature review on hybrid forecasting

Literature review on forecasting problems shows that nowadays research on the application of artificial intelligence methods is developing very dynamically. According to Hall "a new generation of artificial intelligence technologies have emerged that hold considerable promise in helping improve the forecasting process including such applications as product demand, employee turnover, cash flow, distribution requirements, manpower forecasting, and inventory" [29]. Typically manufacturing/distribution planning decisions focus on achieving following goals: "(1) set overall production levels for each product category for each source (manufacturer) to meet fluctuating or uncertain demands for various destinations (distributors) over the intermediate planning horizon and, (2) generate suitable strategies regarding regular and overtime production, subcontracting, inventory, backordering and distribution levels, thereby determining the appropriate resources to be utilized"[46]. Increasing number of decision-making problems related to manufacturing systems (e.g. production scheduling or optimal arrangement of machines) can be solved with the use of algorithms like genetic algorithm, Artificial Bee Colony Algorithm [50] or Tabu Search [13]. There is also a significant number of papers focusing on the application of the artificial intelligence methods for prediction of different aspects related to manufacturing processes (e.g. artificial neural networks (ANN) for predictive compensation of thermal deformations of ball screws in CNC machines [62] or for prediction of average surface roughness and formability [51]) or maintenance in general (e.g. forecasting of mains reliability [77], intelligent forecasting of automatic train protection system failure rate [35], modified convolutional neural network for intelligent fault diagnosis of industrial gearbox [45], ANN-based failure modeling of classes of aircraft engine components [2] or hybrid fault diagnosis of railway switches [54]).

Simultaneously it can be observed that a number of publications on hybrid forecasting is growing rapidly [28]. It is said that hybrid modelling is developed to improve the accuracy of forecasts obtained through the use of individual models. What is more, it is assumed that the forecasts based on combining several methods are simply more accurate than individual ones. According to Hajirahimi and Khashei main advantages of hybrid forecasting models listed in a number of research papers include: "improving forecasting accuracy due to comprehensive pattern detection and modeling", "reducing the risk of using inappropriate model due to the combination of forecasts" and "simplifying the procedure of model selection due to the use of different components"[28].

In [28] Hajirahimi and Khashei performed an in-depth analysis of hybrid forecasting structures on the basis of 150 research papers focused on various hybrid models in time series modeling and forecasting domains. They proposed a classification of hybrid models covering three main combination structures, namely: parallel, series, and parallel-series [28]. This study presents a very detailed and up-to-date review on hybrid forecasting approaches. Nevertheless, taking into account forecasting fields addressed in the analyzed papers, namely [28]: stock market, interest rate, bank circulation, oil price and demand, wind energy, power and speed, tourism and passenger, stream flow, traffic flow, exchange rate, weather and pollutant, GDP, throughput, solar, health, sales and demand, production, hot rolling, internet

Table 1. Review on hybrid forecasting approaches

Hybrid approach	DM	ANN	TQTF	QLF	OM	Implementation area	Source
Combining a multi-layered perceptron neural network and a traditional recursive method		x			x	Spare parts demand forecasting in process industries	[5]
Combining Support Vector Machine (SVM) as a classification tool and autoregressive integrated moving average (ARIMA) model	x		x			Remaining useful life prediction for real-time monitoring of the manufacturing process	[42]
Integrating "the demand autocorrelated process and the relationship between explanatory variables and the nonzero demand of spare parts during forecasting occurrences of nonzero demands over lead times"			x			Spare parts lumpy demand forecasting in the petrochemical industry	[30]
Using SVM model to forecast occurrences of nonzero demand of spare parts and then integrating the forecast being an output from the SVM and the relationship of occurrence of nonzero demand with explanatory variables	x		x			Forecasting intermittent demand of spare parts	[31]
Outputs of moving average (MA) and exponential smoothing as inputs to an ANN model		x	x			Sales forecasting in furniture industry	[66]
Combining Syntetos-Boylan method (being a modification Croston's method) and exponential smoothing			x			Spare parts demand forecasting	[10]
Two-stage approach to forecast intervals of market clearing prices (MCPs) – at first extreme learning machine (ELM) is used to estimate point forecasts of MCPs, next the maximum likelihood method is applied to estimate the noise variance	x				x	Forecasting of electricity prices	[81]
Historical sales data, popularity of article titles, and the prediction result of a time series based on ARIMA are inputs to backpropagation neural network (BPNN)		x	x			Sales forecasting in the publishing industry	[52]
Hybrid of ARIMA model and ANN model		x	x			Short-term price forecasting in deregulated market	[6]
Integrating empirical mode decomposition (EMD), long short-term memory (LSTM) and ELM	x				x	Forecasting of biofuel production	[85]
Hybrid feature selection method (HFS) combining Cuckoo search-based feature selection with singular spectrum analysis and SVM	x				x	Short-term electricity price forecasting	[86]
Combining seasonal autoregressive integrated moving average (SARIMA) model in the ANN model		x	x			Number of inspections forecasting	[67]
Combinations of Kalman filtering (KF), Wavelet Neural Network (WNN) and ANN schemes	x	x				Short-term load forecasting	[3]
Combining adaptive Fourier decomposition, quantitative identification of the average periodicity length and the sine cosine optimization algorithm to select the penalty and kernel parameters of SVM	x		x		x	Electricity demand time series forecasting	[44]
Combining the ARIMA model with time delay neural network (TDNN) and with nonlinear support vector regression (NLSVR) model.	x	x	x			Production forecasting	[60]
Merging the principal component regression method (PCR), the partial least squares regression method (PLSR) and the modified partial least squares regression method (MPLSR).			x		x	Forecasting of product quality evaluation	[84]
Consisting of an ARIMA model and feed-forward, backpropagation network structure with an optimized conjugated training algorithm		x	x			Quality prediction	[53]
Combining Stepwise Regression Method and RBF Neural Network		x	x			Production forecasting	[83]
Combining nonlinear autoregressive with exogenous input (NARX) model and autoregressive moving average (ARMA) model for long-term machine state forecasting based on vibration data.	x		x			Long-term machine state forecasting	[58]
Combining the SARIMA and computational intelligence techniques such as ANN and fuzzy models	x	x	x			Production value of machinery industry forecasting	[36]
Combining traditional forecasting techniques based on time series with artificial intelligence-based methods (ANN and SVM)	x	x	x			Spare parts demand forecasting in mining industry	[64]
Hybrid SVM-based models where three optimization algorithms: gray wolf optimization, whale optimization algorithm and moth flame optimization where applied to optimize the hyper-parameters of the SVM	x				x	Advance rate forecasting of a tunnel boring machine	[87]
Combining a Mahalanobis–Taguchi System (MTS), support vector regression (SVR), bootstrap prediction interval (PI), and derivative-free Nelder-Mead (NM) optimisation strategy.	x				x	Prediction-based multivariate manufacturing process quality control	[72]

traffic, morbidity, birth immigration, NN3 competition, electricity, energy consumption, computer science, rainfall, drought, quality, inflation, solid waste generation, machine state, property crime rates, tidal current, inspection, price, it can be noticed that little research has been done on forecasting dedicated to manufacturing systems. Considering above-listed fields it can be indicated that only a few of them are directly or indirectly related to manufacturing systems – in this aspect forecasting of sales, demand, throughput, production, energy consumption, quality, machine state, inspection and price is significant and should be further investigated.

In general, it can be stated that researchers merge predictive methods and models in very different ways. Table 1 presents an analysis and summary of research results related to hybrid forecasting, which can be applied to specific areas of manufacturing systems. The first column contains a brief description of each hybrid approach, columns 2-6 summarize type of calculation techniques addressed in particular approaches (ANN stands for artificial neural networks, DM stands for data mining techniques (other than ANN), TQTF stands for traditional quantitative forecasting methods, QLF stands for qualitative forecasting methods and OM stands for other methods), whereas the column 7 indicates the implementation area.

According to [28], in general, hybrid forecasting models can be divided into four main groups – data preprocessing based hybrid models, parameters optimization based hybrid models, component combination based hybrid models, and post processing based hybrid models. Analysis of the Table 1 in turn leads to a conclusion that the majority of proposed hybrid methods which can be applied to forecasting manufacturing-related phenomena combines ANN models with traditional quantitative forecasting techniques (especially ARIMA). Quite common is also merging data mining techniques (other than ANN, like e.g. SVM) with ANN models and TQTF. Another noticeable trend is combining DM with other classical mathematical or statistical methods. Very rare is on the other hand combining ANN and qualitative forecasting techniques. "The possibility of generalization of knowledge on new data (that were not presented in the learning process) is an essential characteristic that distinguishes artificial neural networks (ANN)" and thus makes ANN models very often used in hybrid forecasting [63].

Apart from the above-presented analysis which focuses on hybrid forecasting applied to different areas of manufacturing systems, interesting research on hybrid forecasting can be found in results of the M4 Competition, which "follows on from the three previous M competitions, the purpose of which was to learn from empirical evidence both how to improve the forecasting accuracy and how such learning could be used to advance the theory and practice of forecasting" [48]. "The field of forecasting has progressed a great deal since the original M Competition, which concluded that "more complex or statistically sophisticated methods are not necessarily more accurate than simpler methods", and over time, new methods have been proposed that have clearly proven to be more accurate than simpler ones"[48]. From the point of view of the research discussed in this paper the most interesting results following from the M4 Competition include a hybrid and hierarchical forecasting method, which "utilizes a dynamic computational graph neural network system that enables a standard exponential smoothing model to be mixed with advanced long short term memory networks into a common framework" [74] and "a combination-based approach that combines statistical and machine learning techniques" presented in [34].

"Controlling production systems to match supply and demand in an uncertain environment received considerable attention in the manufacturing systems literature"[49], however results of the above-presented literature analysis show that although increasing number of scientific papers is focusing on hybrid approaches, rather little research has been done on hybrid forecasting models dedicated to manufacturing systems. Therefore this paper aims to fill in this gap. The main goal of the study presented in this paper is to propose new artificial intelligence-based hybrid forecasting models and assess their accuracy in comparison to traditional techniques. The research focuses on solving the problem of limited access to explanatory (independent) variables. The research covers three areas of manufacturing, namely: production planning, maintenance and quality control. In order to verify the forecasting accuracy, real data coming from different manufacturing companies are used.

3. Research methodology

Based on the literature review conclusions and bearing in mind experiences gained from cooperation with industrial companies, in this paper new hybrid models are proposed – their goal is to obtain more accurate forecasts in comparison to traditional prediction methods. What is more, the new approach is aiming at solving common problems which still exist in industrial practice (especially in manufacturing companies aiming at transformation into Industry 4.0) – the limited access to data or simply the lack of available data (particularly in terms of explanatory, independent variables). The models are dedicated to forecasting challenges of the manufacturing systems. In the paper four hybrid forecasting models are proposed:

- a Hybrid forecasting econometric model (hybrid_ECO),
- a Hybrid forecasting artificial neural network model (hybrid_ANN),
- a Hybrid forecasting support vector machine model (hybrid_SVM),
- a Hybrid forecasting extreme learning machine model (hybrid_ELM).

The research methodology is schematically presented in Fig. 1. It is composed of 4 main steps: (1) preparation phase, (2) forecasts computation based on traditional forecasting methods, (3) hybrid forecasting models development and (4) assessment phase. The calculations are done in the *R language* (R version 3.5.3), in which a dedicated algorithm was developed.

According to the presented scheme (Fig. 1), in the preparation phase, a forecasting aim and a dependent variable y should be defined. Next, data should be collected – either from appropriate systems (e.g. Enterprise Resource Planning (ERP), Computerised Maintenance Management System (CMMS)) or any adequate database (DB_x), which can contain for instance data coming from the sensors mounted on the machines. Depending on the forecasting aim, the required dataset will differ. Subsequently, the collected data should be initially analyzed and processed. The data is initially divided into a training set (80%) and a test set (20%). In the second phase a parameter should be defined. This parameter on one hand represents the value of a variable which indicates from how many periods an average – in average-based forecasting methods – should be computed and on the other hand it defines how many loops the algorithm will implement. For example for the scope $2 \leq a \leq 3$, the algorithm will compute 2 loops – in the first one, for a=2, all the average-based forecasting methods will apply the average from the last 2 periods, whereas for a=3, in all the average-based forecasting methods the average will be calculated from the last 3 periods. In the next step the 9 analyzed forecasting methods are applied and corresponding forecasts (F_1-F_9) are computed. In the algorithm following methods are implemented:

- F_1: autoregressive-integrated moving average (ARIMA),
- F_2: simple exponential smoothing (SES),
- F_3: Holt model (Holt),
- F_4: trigonometric exponential smoothing (TES),
- F_5: simple moving average (SMA),
- F_6: exponential moving average (EMA),
- F_7: weighted moving average (WMA),
- F_8: zero-lag exponential moving average (ZLEMA),
- F_9: Syntetos-Boylan method (SBA).

Formulas describing each of the 9 analyzed methods are given in Table 2. To check components of all formulas please refer to sources

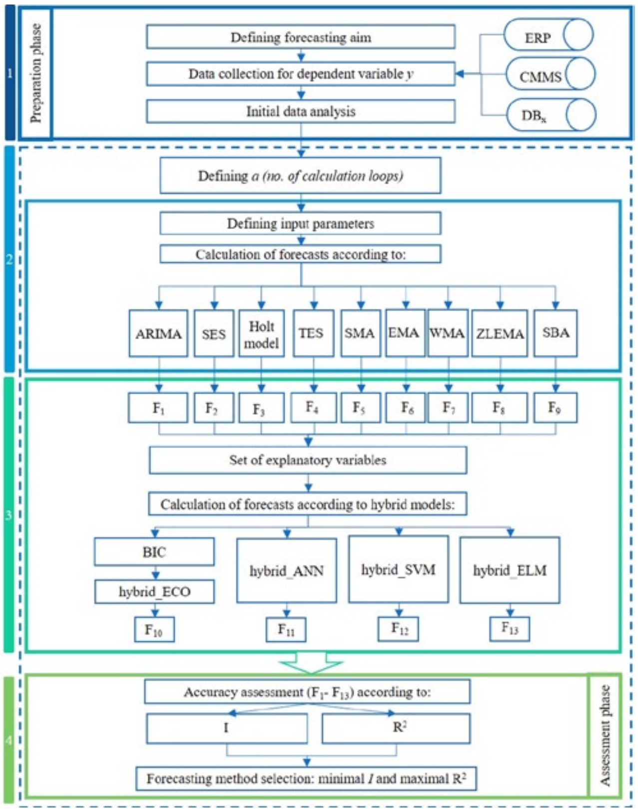


Fig. 1. Algorithm for the forecasting accuracy assessment and optimal method selection

given in the last column ($\hat{y}_t$ is a forecasted value of the variable y in the t period).

In the phase 3, after the forecasts according to 9 traditional methods are calculated, hybrid models are developed – one based on econometric modeling (hybrid_ECO) and three based on artificial intelligence – Hybrid ANN model (hybrid_ANN), Hybrid SVM model (hybrid_SVM) and Hybrid ELM model (hybrid_ELM). Explanatory variables set (EXS) is composed of forecasts coming from the 9 traditional methods (F_1-F_9). It represents an input to each of the proposed hybrid model.

Hybrid econometric model (hybrid_ECO) can be described by the following expression:

$$\hat{y}_t = \hat{\alpha}_0 + \hat{\alpha}_1 F_1 + \hat{\alpha}_2 F_2 + \cdots + \hat{\alpha}_m F_m \tag{10}$$

where: $\hat{\alpha}_i$ – parameters, F_i – explanatory variable composed of the forecasts. For constructing the hybrid_ECO model the Bayesian Schwarz information criterion (BIC) is used to select appropriate subset of explanatory variables from the EXS [1, 68]:

$$BIC = -2\log(L) + p\log(n) \tag{11}$$

where: p – number of model's parameters, L – the maximized value of the likelihood function of the model, n – sample size.

Hybrid ANN model (hybrid_ANN) is developed on the basis of neurons, where the output h_i of neuron i is given by the following formula [69]:

Table 2. Formulas of the forecasting methods

Method	Formula	No./ Source	
ARIMA	$\hat{y}_t = c + \Phi_1 y'_{t-1} + \cdots + \Phi_p y'_{t-p} + \Theta_0 \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \cdots + \Theta_q \varepsilon_{t-q}$	(1)	[33]
SES	$\hat{y}_t = \alpha y_{t-1} + (1-\alpha)\, \hat{y}_{t-1}$	(2)	[33]
Holt	$\hat{y}_t = F_{t-1} + S_{t-1}, \quad F_t = \alpha y_t + (1-\alpha)(F_{t-1} + S_{t-1})$ $S_t = \beta(F_t - F_{t-1}) + (1-\beta) S_{t-1}$	(3)	[33]
TES	$\hat{y}_t = l_{t-1} + \varphi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t \,,\; l_t = l_{t-1} + \varphi b_{t-1} + \alpha d_t \,,\; b_t = (1-\varphi) b + \varphi b_{t-1} + \beta d_t \,,\; s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t$ $d_t = \sum_{i=1}^{p} \rho_i d_{t-1} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-1} + \varepsilon_t$	(4)	[20]
SMA	$\hat{y}_t = \frac{1}{m} \sum_{j=-k}^{k} y_{t+j}$	(5)	[33]
EMA	$\hat{y}_t = \frac{y_{t-1} + (1-\alpha) y_{t-2} + (1-\alpha)^2 y_{t-3} + \ldots + (1-\alpha)^n y_{t-(n+1)}}{1 + (1-\alpha) + (1-\alpha)^2 + (1-\alpha)^3 + \ldots + (1-\alpha)^n}$	(6)	[59]
WMA	$\hat{y}_t = \sum_{j=-k}^{k} a_j y_{t+j}$	(7)	[33]
ZLEMA	$\hat{y}_t = \frac{2}{(n+1)}(2y_{t-1} - y_{lag}) + \left(1 - \frac{2}{(n+1)}\right) \times \hat{y}_{t-1}$	(8)	[89]
SBA	$\hat{y}_t = \left(1 - \frac{\alpha}{2}\right) \frac{Z_{t-1}}{P_{t-1}}, Z_t = \alpha X_t + (1-\alpha) Z_{t-1}$ $P_t = \alpha G_t + (1-\alpha) P_{t-1}$	(9)	[18]

$$h_i = \sigma\left(\sum_{j=1}^{N} W_{ij}x_i + T_i^{hidden}\right) \quad (12)$$

where: σ() – the transfer function, N – the number of input neurons; W_{ij} – the weights; x_j – inputs to the input neurons, T_i^{hidden} – the threshold of the hidden neurons. More details on ANN development can be found in [21].

Hybrid SVM model (hybrid_SVM) is based on the functional dependence of the dependent variable y on a set of explanatory variables x. The relationship between the explanatory variables and $\hat{y}_t$ is given by a deterministic function f and the addition of some noise [90]:

$$\hat{y}_t = f(x) + noise \quad (13)$$

where x is a set of explanatory variables (x=F_1, F_2, … , F_{12}). The functional form for f which can correctly predict new cases can be achieved by training the SVM model on a sample set – a process involving the sequential optimization of an error function (for details see [90]). Radial basis function (RBF) will be the kernel type K used in the Hybrid SVM model [90]:

$$K(X_i, X_j) = \exp(-\gamma |X_i - X_j|^2) \quad (14)$$

where $K(X_i, X_j) = \phi(X_i) \cdot \phi(X_j)$, ϕ – transformation.

Hybrid extreme learning machine model is based on the algorithm which can be summarized as follows [32]: "given a training set $\aleph = \{(x_i, t_i) | x_i \in \boldsymbol{R}^n, t_i \in \boldsymbol{R}^m, i = \ldots, N\}$, activation function $g(x)$, and hidden node number $\tilde{N}$,

Step 1: Randomly assign input weight w_i and bias b_i , i=1, …, $\tilde{N}$.
Step 2: Calculate the hidden layer output matrix **H**.
Step 3: Calculate the output weight β

$$\beta = H \; T \quad (15)$$

where $\boldsymbol{T} = [t_1, \ldots, t_N]^T$".

After the hybrid models are developed, the last step of the research methodology can be applied – accuracy assessment of estimated forecasts for each analyzed method.

The third phase is finished when 4 forecasts (F_{10} - F_{13}) from hybrid models are computed. The last, fourth phase, is the assessment phase. The algorithm allows to compute five types of forecasts accuracy measures, namely: mean error (ME), mean absolute error (MAE), root mean squared error (RMSE), relative forecast error *ex post* (I) and coefficient of determination (R^2), yet in the proposed methodology, the selection of the most accurate forecasting method is based on the value of I. Therefore, the best model is the one with the lowest I. Relative forecast error *ex post* I is given by the formula:

$$I = \sqrt{\frac{\sum_{t=1}^{m}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{m} y_t^2}} \quad (16)$$

In case of models with equal I, as the most accurate will be considered the one with the highest coefficient of determination R^2 , which can be defined as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_t - \bar{y})^2}{\sum_{i=1}^{n}(y_t - \bar{y})^2} \quad (17)$$

According to the presented approach, the calculations are over when the algorithm computes all the accuracy measures in all loops for the 13 considered methods (9 traditional and 4 hybrid) and – according to the above-mentioned rule – indicates the most effective method for the given forecasting aim. The proposed algorithm developed in *R language* is very flexible – it can be easily adjusted by adding other forecasting methods or – if necessary – more accuracy measures. It can serve as a supporting tool in the decision-making process of manufacturing companies trying to select the most appropriate forecasting method. It can also be considered as an approach supporting the transformation process of Industry 4.0 implementation in industrial factories. According to Bożejko et al., nowadays, in the ERP systems supporting the management, "particularly important are numerically efficient methods and algorithms for solving the new optimization problems derived from real manufacturing systems which constitute "intelligent engines" for these support systems" [11]. What is more, they are considered as critical for the production efficiency of large manufacturing companies. The proposed algorithm can be integrated into such support systems.

The further research discussed in this paper will be carried out according to the methodology presented in this chapter. It will be applied to 3 case studies and the calculations will be performed on real data from the manufacturing companies. In order to exemplify manufacturing areas, which will be considered in the practical part of this paper, the model of the manufacturing system was developed (Fig. 2).

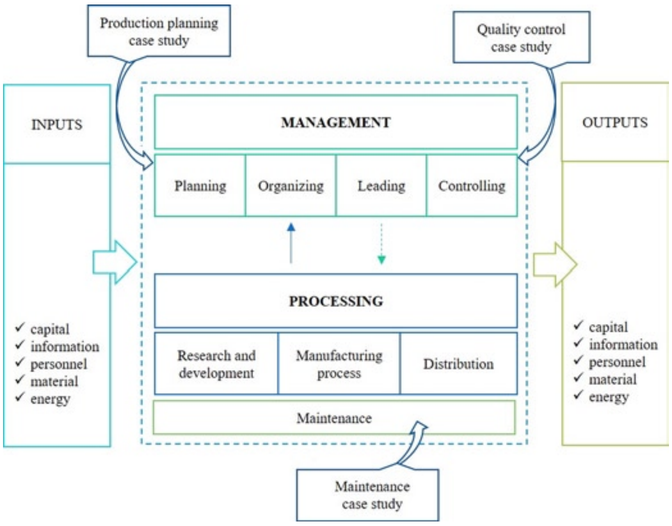


Fig. 2. Manufacturing system model

According to Caggiano a manufacturing (production) system can be defined as "an organization in the manufacturing industry for the creation of production. In the mechanical and electrical engineering industries, a manufacturing system, in general, has an integrated group of functions, e.g., the sales, design, production, and shipping functions"[14]. The case studies presented in the next chapter address three areas of the manufacturing system – production planning, maintenance and quality control. They are schematically presented in Fig. 2.

4. Performance assessment of the hybrid models – real data analysis

4.1. Forecasting in manufacturing systems – case studies

As presented in the Introduction, forecasting in manufacturing systems plays a very important role. In this chapter three case studies referring to particular areas of the manufacturing system will be addressed – production planning, maintenance, and quality control. The Author's intention was to apply the hybrid models to datasets coming not only from different areas of the production system, but also from different manufacturing sectors, therefore to verify the proposed models 3 different companies were selected and addressed. The idea behind this approach was to check if the developed algorithms are versatile and comprehensive enough to meet forecasting challenges from various industries.

4.1.1. Production planning case study

The first case study is related to the forecasting of a product demand in a furniture factory. The details of the research concerning this case can be found in [66]. The main problem in the addressed company, which sells its products mainly via Internet, was to develop a new manufacturing system that would allow to increase effectiveness and production volume, reduce delivery time to 48 hours (in the online sales channel) and to compute more accurate sales forecasts, which, in turn, would result in more efficient production planning. Development of the manufacturing system answering all these challenges was a demanding and a complex task. While constructing the concept of the expected system, several research questions were raised, namely: (1) how the sales level can be predicted, if the company did not have a sufficient and reliable database; (2) how in such circumstances a set of explanatory variables can be prepared, (3) how to control the manufacturing process if the demand was very diversified – series production and individual, customized orders; (4) how to develop production strategy for various product types; and last but not least, (5) how to verify, if the proposed system was immune to disturbances. The answers to these questions were addressed in [66]. In this paper however, only selected aspects will be tackled – particularly – how to select the right forecasting method and how to assure the most accurate forecasts to given datasets. Due to the reason that "future demand plays a very important role in production planning and inventory management, fairly accurate forecasts are needed" [26]. What is more, "in the customization processes, it is important to keep the manufacturing system reliable, therefore, a prognostic method is essential" [73]. For the purpose of this study 5 datasets were investigated – the basic information about them contains Table 3. Moreover, for each product a box plot (Fig. 3) and a histogram (Fig. 4) were developed.

Table 3. Production planning case study – summary of the datasets

No.	Product	Unit	n	Demand type	Expected result: Forecasted value of variable y
1	Product A	piece	23	monthly	Expected monthly demand for product A (e.g. 10 pieces)
2	Product B	piece	23	monthly	Expected monthly demand for product B (e.g. 20 pieces)
3	Product C	piece	23	monthly	Expected monthly demand for product C (e.g. 30 pieces)
4	Product D	piece	23	monthly	Expected monthly demand for product D (e.g. 15 pieces)
5	Product E	piece	23	monthly	Expected monthly demand for product E (e.g. 5 pieces)

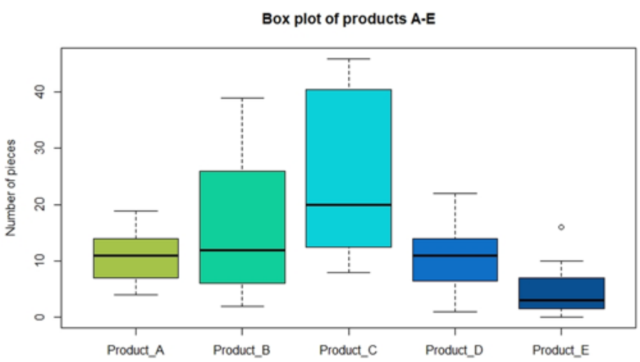


Fig. 3. Box plots of products A-E

In order to verify the accuracy of the proposed in this paper hybrid forecasting models and compare their effectiveness with other methods, 5 types of products were investigated (Products A-E). Due

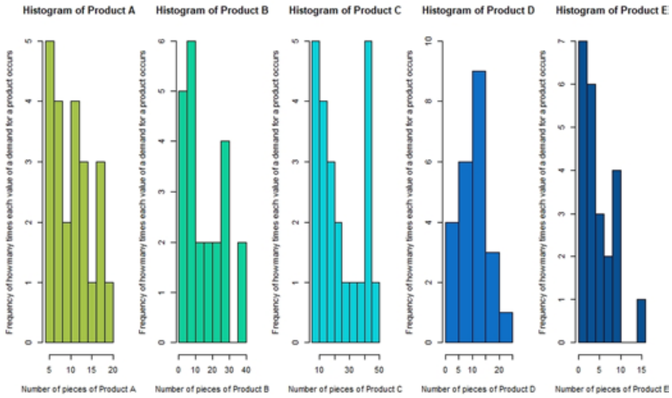


Fig. 4. Histograms of products A-E

to unsatisfactory precisions of daily forecasts of demand for selected types of furniture (caused by a shortage of available historical data in the investigated company), the analysis was carried-out in aggregated, monthly terms.

4.1.2. Maintenance case study

The second case study is related to the forecasting of spare parts and consumable materials demand in a copper mine. The details of the research concerning this case can be found in [16, 64, 65]. This study is related to the aspects of preventive maintenance, which can be defined as: "maintenance executed at predetermined intervals or according to prescribed criteria, aiming to reduce the probability of failure or the probability that an item will only fulfill its functions to a limited extent (degradation of functioning)", where an item is "any part, component, devise, subsystem, functional unit, equipment, or system that can be individually considered. Failure is regarded as the termination of the ability of an item to perform an action as required" [71]. Predictive maintenance "helps to avoid downtimes due to unexpected failures during the production process" [22]. When investigating different sectors of economy, it can be observed that in particular enterprises from underground mining are characterized by very high failure rates of machines [43]. The main reasons for that are the very specific working environment characterized by high temperatures, high humidity and poor road conditions. What is more, mining machines are almost constantly in motion. Besides, the complexity of these machines and the high loads to which they are subjected impose very strict requirements on their reliability and maintenance [12, 27]. Mining processes are unstable and cost-intensive, which makes that controlling them very difficult [82]. Hence the aspect of accurate spare parts demand forecasting is essential because it directly influences the availability of machines and their maintenance processes. "When a critical part is requested and not available in stock, the company is not able to perform the maintenance operation in time. This could jeopardize a client's productivity, causing time delays and high costs" [78]. It is also worth mentioning that according to Chen et al. "in practice, the forecast and inventory planning of service parts depend on accurate predictions of product failure rates" [15].

For the purpose of this paper real data from the underground copper mine were used. The data were gathered within the research project "Adaptation and Implementation of Lean Methodology in Copper Mines" co-financed by the Polish National Centre for Research and Development. To assess the forecasting accuracy of the proposed in this study hybrid models in maintenance area 4 datasets were selected, namely: demand for brake pump, actuator, hydraulic oil and diesel

Table 4. Maintenance case study – summary of the datasets

No.	Spare part/ consumable material	Unit	n	Demand type	Expected result: Forecasted value of variable y
1	Brake pump	piece	47	weekly	Expected weekly demand for brake pump (e.g. 5 pieces)
2	Actuator	piece	52	weekly	Expected weekly demand for actuator (e.g. 3 pieces)
3	Hydraulic oil	litre	248	weekly	Expected weekly demand for hydraulic oil (e.g. 500 litres)
4	Diesel oil	litre	313	weekly	Expected weekly demand for diesel oil (e.g. 5000 litres)

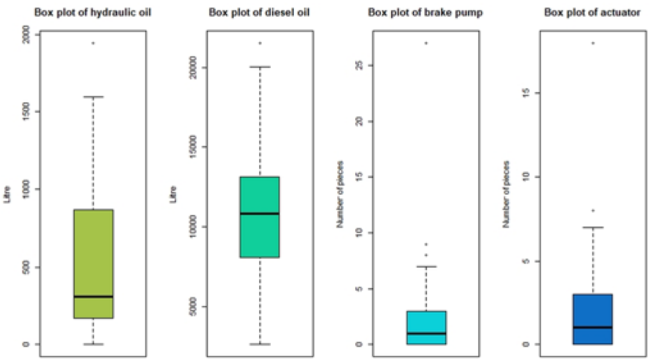


Fig. 5. Box plots of investigated spare parts and consumable materials

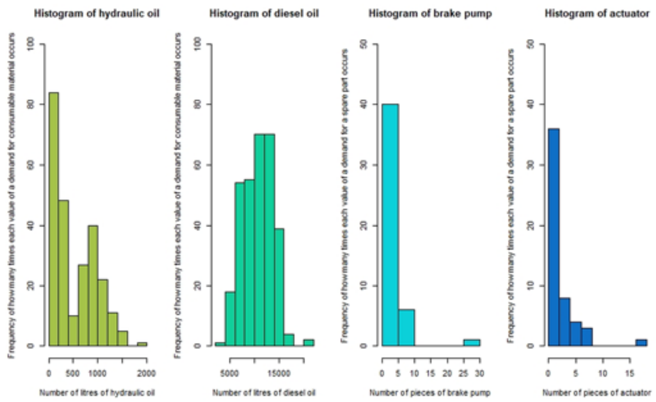


Fig. 6. Histograms of investigated spare parts and consumable materials

Table 5. Quality control case study – summary of the datasets

No.	Dependent variable:	Unit	n	Level of defects	Expected result: Forecasted value of variable *y*
1	level of defects	%	21	monthly	Expected monthly level of defects (e.g. 2%)
2	level of defects	%	22	monthly	Expected monthly level of defects (e.g. 4%)

oil. The basic information about the investigated cases is presented in Table 4, Fig. 5 (box plots) and Fig. 6 (histograms).

Spare parts demand is very hard to predict – it is characterized by the large degree of uncertainty and by unpredictable fluctuations. This type of demand is often classified as lumpy, which can be defined as "a demand with great differences between each period's requirements and with a great number of periods with zero requests" [23] or as intermittent, which means that this demand "is characterised by variable demand sizes coupled with irregular demand arrivals, with many observations having zero demand"[57]. An in-depth literature review on the spare parts demand forecasting can be found in works by Bacchetti and Saccani, who discuss spare parts classification and demand forecasting for stock control [7], Van Horenbeek et al., who present a review on joint maintenance and inventory optimization systems [79], Rego and Mesquita, who have developed a literature review on spare parts inventory control [61], and De Gooijer and Hyndman, who have investigated and described 25 years of time series forecasting [19]. An accurate forecasting of spare parts demand is very challenging – usually obtained predictions are characterized by large errors. Therefore, forecasting in maintenance area is especially hard, yet at the same time very important. It is worth mentioning that the management of spare parts is considered to be of the most neglected areas of management whereas its meaning cannot be overemphasized [24, 65].

4.1.3. Quality control case study

The third case study addressed in this paper was focused on the forecasting of defects. The data which was used in order to assess the forecasting accuracy of the proposed in this paper hybrid models, came from an industrial company which manufactured ceramic insulators. In this example the main challenge was to develop an efficient solution supporting process control in production of ceramic insulators to ensure the desired product quality. The details of this study can be found in [40]. The goal of this research was to find a correlation between grain-size distribution of aluminum oxide and the number of quality defects. It was assumed that it was possible to control addition of raw aluminum oxide (and its graining) to obtain its desired grain-size composition in the mass and thus to reduce to acceptable level the number of insulators' defects, namely: (1) cracks (on bodies, on sheds, on face surfaces and in holes), (2) twists and (3) disturbed structure. In this research 2 datasets from 2 different periods were investigated – their summary is presented in Table 5. What is more, for each dataset a box plot (Fig. 7) and a histogram (Fig. 8) were developed.

In this case study the forecasted variable y is an expected monthly level of defects. The challenge in this case was to create effective models (model A based on the first data set, model B based on the second data set) enabling accurate forecasting of the level of defects. For the purpose of this study the explanatory variables were ignored and only dependent variable was investigated. It is expected that the analyzed level of defects should be easily forecasted and controlled. It should be underlined that the precise defects forecasting is one of the pillars of the effective quality control.

4.2. Accuracy assessment of the hybrid models

According to the research methodology presented in the previous chapter, computations were carried out in *R language*. The scope of *a* parameter was set from 2 to 7, which means that the algorithm computed 6 loops for each dataset. In total 11 datasets were investigated – 5 products A-E from production planning case study, 4 datasets from maintenance case study (brake pump, actuator, hydraulic oil, diesel oil) and 2 datasets from quality control case study. The accuracy as-

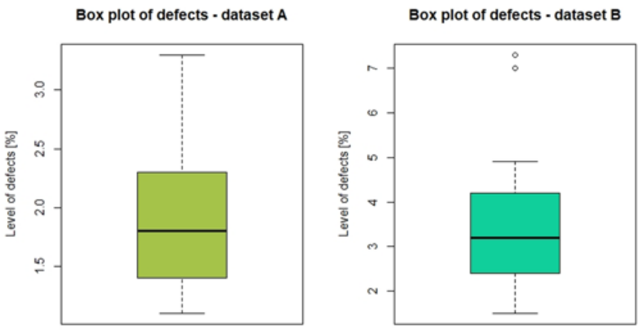


Fig. 7. Box plots presenting levels of defects [%] – datasets A and B

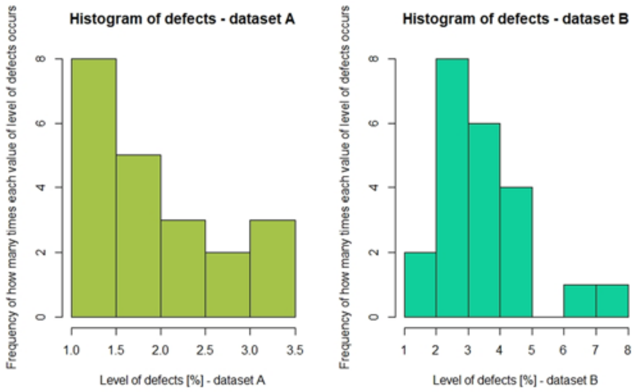


Fig. 8. Histograms presenting levels of defects – datasets A and B

*Table 6. Relative forecast error **ex post** for production planning case study*

Method	Production planning				
	Product A	Product B	Product C	Product D	Product E
	a=4	a=2	a=5	a=6	a=7
ARIMA	0,374	0,271	0,399	0,353	0,421
SES	0,504	0,281	0,445	0,379	0,458
Holt	0,433	0,282	0,445	0,365	0,422
TES	0,343	0,270	0,437	0,338	0,311
SMA	0,385	0,258	0,466	0,360	0,535
EMA	0,414	0,272	0,436	0,369	0,490
WMA	0,408	0,262	0,440	0,352	0,472
ZLEMA	0,486	0,322	0,443	0,377	0,419
SBA	0,608	0,611	0,489	0,443	0,600
hybrid_ECO	**0,118**	0,226	0,292	0,242	**0,155**
hybrid_ANN	0,446	**0,161**	**0,142**	**0,167**	0,316
hybrid_SVM	0,289	0,206	0,258	0,271	0,276
hybrid_ELM	0,216	0,193	0,303	0,219	0,165

*Table 7. Relative forecast error **ex post** for maintenance case study*

Method	Maintanance			
	Brake pump	Actuator	Hydraulic oil	Diesel oil
	a=6	a=7	a=2	a=6
ARIMA	0,887	0,819	0,282	0,183
SES	1,160	1,052	0,315	0,211
Holt	1,061	0,971	0,298	0,207
TES	0,825	0,821	0,282	0,182
SMA	0,949	0,886	0,302	0,201
EMA	0,965	0,878	0,301	0,194
WMA	0,971	0,892	0,306	0,199
ZLEMA	1,099	0,994	0,407	0,222
SBA	0,900	1,301	0,539	0,197
hybrid_ECO	0,815	0,768	0,284	0,180
hybrid_ANN	**0,260**	**0,550**	**0,263**	**0,157**
hybrid_SVM	0,635	0,638	0,266	0,174
hybrid_ELM	0,731	0,744	0,280	0,180

sessment was performed in two steps – at first analysis based on relative forecast error *ex post I* was done, and secondly analysis based on coefficient of determination R^2.

For each case study a dedicated table was prepared covering values of relative forecast error *ex post I* calculated for every analyzed method. Table 6 contains obtained results for production planning case study – particular columns represent products A-E and corresponding values of *a* parameter for which the highest accuracy was obtained (the lowest value of *I* for the most accurate method). In rows all 13 investigated methods are listed. The lowest value of *I* for the most accurate method for each product is marked in bold.

Analysis of obtained results in Table 6 shows that in case of 3 products out of 5 the most accurate method is hybrid_ANN. For the products A and E the lowest *I* were obtained by Hybrid forecasting econometric model (hybrid_ECO). In general it can be said that for production planning the forecasts from the most efficient hybrid models are very accurate – relative forecast errors *ex post* vary from 11,8% to 16,7%.

Table 7 presents computation results for the maintenance case study. There were 4 types of datasets analyzed – 2 representing spare parts (brake pump, actuator) and 2 representing consumable materials (hydraulic oil, diesel oil).

The values of a parameter for which the highest accuracy was obtained for particular cases are following: a=6 for brake

Table 8. Relative forecast error ex post for quality control case study

Method	Quality control	
	No. of defects A	No. of defects B
	a=4	a=3
ARIMA	0,300	0,333
SES	0,288	0,314
Holt	0,300	0,308
TES	0,292	0,265
SMA	0,327	0,319
EMA	0,292	0,299
WMA	0,300	0,309
ZLEMA	0,323	0,326
SBA	0,334	0,367
hybrid_ECO	0,242	0,168
hybrid_ANN	**0,146**	**0,151**
hybrid_SVM	0,217	0,260
hybrid_ELM	0,251	0,223

pump, a=7 for actuator, a=2 for hydraulic oil and a=6 for diesel oil. In maintenance case study the lowest values of relative forecast errors ex post were obtained by hybrid_ANN models, which delivered the most accurate forecasts. Analogous to the production planning and the maintenance case studies, the calculations were performed in terms of the quality control example. The obtained results are presented in Table 8. In this research only 2 cases were investigated – level of defects based on 2 datasets (A and B).

For both analyzed datasets, the most accurate method of forecasting level of defects turned out to be Hybrid forecasting ANN model (I = 14,6% for dataset A, I = 15,1% for dataset B). Other investigated methods were less accurate.

In order to assess the accuracy of the analyzed forecasting methods in terms of particular manufacturing system's areas, an average relative forecast error *ex post* was computed in percentage terms for each case study separately. Table 9 contains the obtained results. The average I values were divided into three groups:

- accurate forecasts: $I \leq 25\%$ (results marked **in bold**);
- moderately accurate forecasts: 25% <I <50% (results marked in light grey);
- not accurate forecasts: $I \geq 50\%$ (results marked in dark grey).

*Table 9. Comparison of the forecasting methods accuracy based on **I***

Method	Average relative forecast error ex post		
	Production planning	Maintanance	Quality control
ARIMA	36%	54%	32%
SES	41%	68%	30%
Holt	39%	63%	30%
TES	34%	53%	28%
SMA	40%	58%	32%
EMA	40%	58%	30%
WMA	39%	59%	30%
ZLEMA	41%	68%	32%
SBA	55%	73%	35%
hybrid_ECO	**21%**	51%	**21%**
hybrid_ANN	**25%**	31%	**15%**
hybrid_SVM	26%	43%	**24%**
hybrid_ELM	**22%**	48%	**24%**

An analysis of obtained results leads to several conclusions. First of all it can be noticed that only in 2 manufacturing areas accurate forecasts were obtained – in production planning and in quality control. In case of maintenance all methods provided either moderately accurate forecasts (for the proposed artificial intelligence based hybrid models: $31\% \leq I \leq 48\%$) or not accurate forecasts at all (all the other 10 investigated methods). In production planning only Hybrid forecasting econometric model (average I = 21%), Hybrid forecasting ANN model (average I = 25%) and Hybrid forecasting ELM model (average I = 22%) provided satisfactory results – the obtained I was not higher than 25%. Among conventional forecasting methods SBA turned out to be not accurate, whereas the other methods can be classified as moderately accurate. In quality control all the four proposed hybrid models (hybrid_ECO, hybrid_ANN, hybrid_SVM, hybrid_ELM) provided accurate forecasts (average I did not exceed 24%). Other analyzed methods – ARIMA, SES, Holt, TES, SMA, EMA, WMA, ZLEMA, SBA – occurred to generate moderately accurate forecasts.

In order to assess the accuracy of the four proposed hybrid models in comparison to 9 other researched forecasting methods an average I was computed (in percentage terms) and presented in Fig. 9.

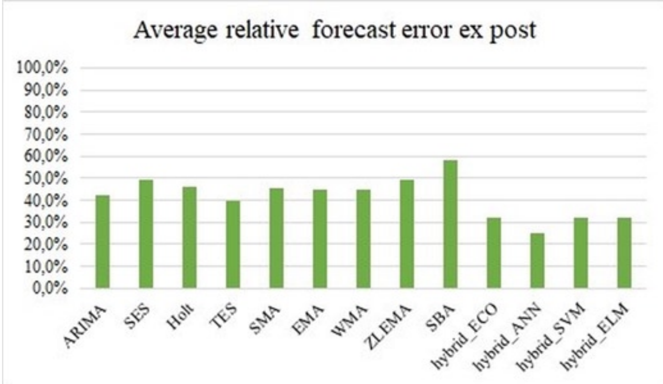


Fig. 9. Average relative forecast error ex post for analyzed methods

The average I was computed from 11 investigated datasets for each method separately. Obtained results show that the proposed in this paper hybrid models deliver much more precise forecasts in comparison to other researched methods. The accuracy of estimated forecasts from hybrid models is significantly higher. An analysis of efficiency of other methods leads to a conclusion that for investigated cases they seem to be not effective.

The second part of the accuracy assessment was based on R^2. The analysis was performed analogous to the above presented. For each case study a dedicated table was prepared with values of coefficient of determination computed for every analyzed method. Table 10 contains obtained results for production planning case study – particular columns represent products A-E and corresponding values of *a* parameter for which the highest accuracy was obtained (the highest value of R^2 for the most accurate method). In rows all 13 investigated methods are listed. The highest value of R^2 for the most accurate method for each product is marked in bold.

Analysis of obtained results shows that in case of 3 products out of 5 the most accurate method is Hybrid forecasting artificial neural network – R^2 in percentage terms is equal respectively for product B 93,0%, for product C 88,2%, for product D 77,1%. For product A and E the highest R^2 was obtained by hybrid_ECO models (R^2 = 91,3%, R^2 = 92,7%). These results confirm the conclusion derived from

Table 10. Coefficient of determination for production planning case study

Method	Production planning				
	Product A	Product B	Product C	Product D	Product E
	a=4	a=2	a=5	a=6	a=7
ARIMA	0,278	0,837	0,110	0,005	0,520
SES	0,024	0,820	0,119	0,002	0,447
Holt	0,087	0,823	0,118	0,006	0,486
TES	0,306	0,824	0,106	0,072	0,740
SMA	0,207	0,852	0,045	0,018	0,410
EMA	0,125	0,819	0,075	0,008	0,496
WMA	0,143	0,844	0,086	0,009	0,488
ZLEMA	0,042	0,737	0,185	0,016	0,537
SBA	0,232	0,452	0,014	0,002	0,498
hybrid_ECO	**0,913**	0,852	0,484	0,373	**0,927**
hybrid_ANN	0,406	**0,930**	**0,882**	**0,771**	0,710
hybrid_SVM	0,555	0,885	0,633	0,272	0,836
hybrid_ELM	0,710	0,892	0,446	0,487	0,917

I analysis – it can be stated that for production planning the forecasts from hybrid models are accurate – in case of all investigated products R^2 exceeded 77%.

Table 11 presents computation results for maintenance case study, in which 4 types of datasets were analyzed.

Table 11. Coefficient of determination for maintenance case study

Method	Maintanance			
	Brake pump	Actuator	Hydraulic oil	Diesel oil
	a=6	a=7	a=2	a=6
ARIMA	0,231	0,561	0,794	0,549
SES	0,001	0,000	0,753	0,461
Holt	0,007	0,001	0,776	0,463
TES	0,156	0,020	0,794	0,562
SMA	0,039	0,001	0,768	0,469
EMA	0,021	0,000	0,770	0,502
WMA	0,022	0,000	0,764	0,482
ZLEMA	0,010	0,003	0,639	0,417
SBA	0,011	0,022	0,618	0,489
hybrid_ECO	0,156	0,121	0,788	0,562
hybrid_ANN	**0,914**	0,573	**0,820**	**0,668**
hybrid_SVM	0,639	**0,743**	0,817	0,592
hybrid_ELM	0,321	0,175	0,796	0,567

In maintenance case study the highest values of R^2 were again obtained by hybrid models – in 3 out of 4 cases Hybrid forecasting ANN model delivered the most accurate forecasts – R^2 in percentage terms reached respectively: 91,4% for brake pump, 82,0% for hydraulic oil, and 66,8% for diesel oil. In the case of the actuator the hybrid_SVM model turned out to be the most effective forecasting method (R^2 = 74,3%).

Next, the calculations for quality control example were applied. The obtained results are presented in Table 12. As mentioned before, in this research only 2 cases were investigated – level of defects based on dataset A and level of defects based on dataset B.

Table 12. Coefficient of determination for quality control case study

Method	Quality control	
	No. of defects A	No. of defects B
	a=4	a=3
ARIMA	0,337	0,343
SES	0,342	0,379
Holt	0,283	0,384
TES	0,292	0,637
SMA	0,168	0,401
EMA	0,280	0,493
WMA	0,262	0,415
ZLEMA	0,354	0,418
SBA	0,155	0,540
hybrid_ECO	0,455	0,814
hybrid_ANN	**0,808**	**0,851**
hybrid_SVM	0,579	0,678
hybrid_ELM	0,416	0,674

The obtained values of R^2 confirm the results based on *I* analysis – for both datasets the most accurate method of forecasting level of defects turned out to be Hybrid forecasting ANN model (R^2 = 80,8% for dataset A and R^2 = 85,1% for dataset B).

Assessment of the accuracy of the analyzed forecasting methods in terms of particular manufacturing system's areas, was based on average R^2 (in percentage terms) calculated for each case study separately. Table 13 contains the obtained results. The average R^2 values were divided into three groups:

- accurate forecasts: R^2 >70% (results marked **in bold**);
- moderately accurate forecasts: 61% ≤ R^2 ≤ 70% (results marked in light grey);
- not accurate forecasts: R^2 ≤ 60% (results marked in dark grey).

Table 13. Comparison of the forecasting methods accuracy based on average R^2

Method	Average coefficient of determination		
	Production planning	Maintanance	Quality control
ARIMA	35%	53%	34%
SES	28%	30%	36%
Holt	30%	31%	33%
TES	41%	38%	46%
SMA	31%	32%	28%
EMA	30%	32%	39%
WMA	31%	32%	34%
ZLEMA	30%	27%	39%
SBA	24%	29%	35%
hybrid_ECO	**71%**	41%	63%
hybrid_ANN	**74%**	**74%**	**83%**
hybrid_SVM	64%	70%	63%
hybrid_ELM	69%	46%	54%

An analysis of average R^2 values shows that accurate forecasts were obtained in each analyzed manufacturing system area: in production planning (hybrid_ECO: R^2 = 71%, hybrid_ANN: R^2 = 74%), in maintenance (hybrid_ANN: R^2 = 74%) and in quality control (hybrid_ANN: R^2 = 83%). In case of maintenance example – apart from hybrid_ANN and hybrid_SVM (moderately accurate forecasts) all the other 11 investigated methods provided not accurate forecasts. In production planning hybrid_SVM and hybrid_ELM delivered moderately accurate forecasts, whereas all the other methods can be assessed as not accurate. In quality control, apart from hybrid_ANN, which turned out to provide accurate forecasts, two methods can be classified as moderately accurate (hybrid_ECO and hybrid_SVM), whereas other investigated methods – ARIMA, SES, Holt, TES, SMA, EMA, WMA, ZLEMA, SBA – brought not satisfactory results. On the basis of these results it can be concluded that hybrid models can be more efficient forecasting tools in the areas of production planning and quality control, than in maintenance, where the accurateness of the proposed hybrid methods is lower.

In order to assess the accuracy of the four proposed hybrid models in comparison to 9 other researched forecasting methods an average R^2 was computed (in percentage terms). The average R^2 was calculated from 11 investigated datasets for each method separately (Fig. 10).

The obtained values of average R^2 show that the proposed hybrid models provide significantly more accurate forecasts in comparison to the other 9 researched methods. The average R^2 in percentage terms in case of hybrid_ECO equals 59%, in case of hybrid_ANN 76%, in case of hybrid_SVM 66% and in case of hybrid_ELM 58%, whereas

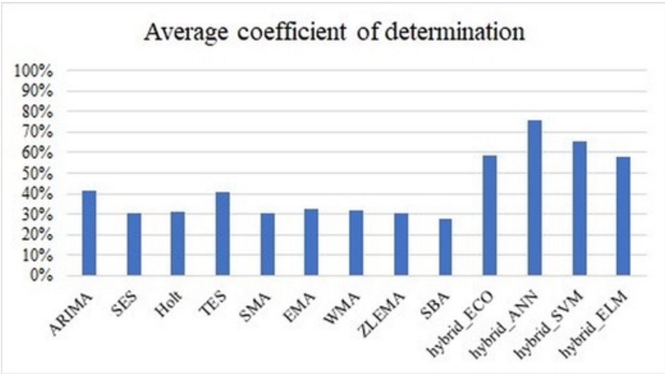


Fig. 10. Average coefficient of determination for analyzed methods

for the other methods R^2 reached only 41% (ARIMA, TES) or less. The least efficient turned out to be conventional time series forecasting methods, namely: SES, Holt, SMA, EMA, WMA, ZLEMA and SBA – R^2 did not reach 34%.

5. Conclusions

The main aim of the paper was to propose the new artificial intelligence-based hybrid forecasting models and assess their accuracy in comparison to traditional techniques. The analysis was performed based on the assumption that the access to explanatory (independent) variables was not possible – lack of corresponding data. The results of the study were satisfactory – the analysis of the forecasting accuracy of the new hybrid models (hybrid_ECO, hybrid_ANN, hybrid_SVM, hybrid_ELM) showed that they are more precise than other investigated methods, namely: ARIMA, SES, Holt, TES, SMA, EMA, WMA, ZLEMA and SBA. Obtained values of the relative forecast errors *ex post I* and the coefficients of determination R^2 proved that hybrid models proposed in this paper are significantly more accurate than the rest of the methods. Moreover, the study fills in the literature gap on application of hybrid forecasting in manufacturing systems. According to the presented research methodology, the proposed models were verified on real data from the three areas of the manufacturing system – production planning, maintenance and quality control. The investigated case studies showed that the proposed hybrid models can serve as efficient forecasting tools in manufacturing companies. The obtained forecasting results were especially satisfactory in terms of production planning and quality control. The accuracy of predictions in maintenance was acceptable, yet less efficient than in two other investigated areas of the manufacturing system. Bearing in mind, however, that the analyzed demand was lumpy and intermittent, the obtained results were sufficient. In general conventional time series forecasting methods were ineffective in the researched areas of the manufacturing system.

Results of the literature review showed also that although an increasing number of scientific papers is focusing on the development of the hybrid forecasting models, the majority of them is combining only a few methods (usually two or three). The distinguishing feature of the proposed hybrid models is that each of them combines in total 10 methods. This approach also helps to solve a common problem in manufacturing companies which is related to the limited access to appropriate data. The preparation of the right set of potential explanatory variables is sometimes impossible due to the lack of available, reliable, complete and comparable statistical data. Due to this reason the companies cannot use the forecasting methods based on the independent variables. The proposed hybrid models solve this problem.

What is more, in the paper, the algorithm for the forecasting accuracy assessment and optimal method selection was introduced. It can serve not only as an efficient and costless tool for advanced manufacturing companies willing to select the right forecasting method for their particular needs, but also as an approach supporting implementation of Industry 4.0 technologies and transformation towards smart factories. It is an important and required solution as still many manufacturing companies are facing the challenge of transformation from the so-called 3rd to the 4th industrial revolution.

The presented case studies showed that the accurate forecasts can efficiently support production planning, quality control and maintenance management, through: (1) controlling the product quality parameters, (2) making a manufacturing process more flexible, (3) reducing failures and (4) elimination of root causes of defects. In consequence, thanks to improved forecasts, the manufacturing companies can reduce their inventory, increase inventory turns and improve order indicators, which brings significant savings and leads to lower costs, increased revenues and thus influences profitability.

The future research will focus on the further development of the algorithm for the forecasting accuracy assessment and optimal method selection – it is planned to add more prediction methods. The proposed hybrid models will be tested on more real datasets from a wider range of manufacturing applications. It is also planned to optimize the models parameters, so that the obtained forecasts are as accurate as possible. Further works include also studies on the implementation of the algorithm to a comprehensive information system which will be an extension of integrated systems currently used in companies (e.g. the Enterprise Resources Planning) or as a part of the company integrated management system. What is more, it is planned to develop the integration capabilities to form a connection with a range of sensors and monitoring equipment so as to collect more accurate machine/device data directly to the algorithm.

References

1. Acquah D-G H. The effect of outliers on the performance of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in selection of an asymmetric price relationship. Russian Journal of Agricultural and Socio-Economic Sciences 2017; 65(5): 32-37, https://doi.org/10.18551/rjoas.2017-05.05.
2. Al-Garni A, Abdelrahman W, Abdallah A. ANN-based failure modeling of classes of aircraft engine components using radial basis functions. Ekspolatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 311-317, https://doi.org/10.17531/ein.2019.2.16.
3. Aly HHH. A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid. Electric Power Systems Research 2020; 182: 106191, https://doi.org/10.1016/j.epsr.2019.106191.
4. Ambrożkiewicz B, Syta A, Meier N et al. Radial internal clearance analysis in ball bearings. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(1): 42-54, https://doi.org/10.17531/ein.2021.1.5.
5. Amin-Naseri MR, Tabar BR. Neural network approach to lumpy demand forecasting for spare parts in process industries. 2008 International Conference on Computer and Communication Engineering, Kuala Lumpur, Malaysia, IEEE: 2008: 1378-1382, https://doi.org/10.1109/ICCCE.2008.4580831.
6. Areekul P, Senjyu T, Toyama H, Yona A. Notice of Violation of IEEE Publication Principles - A Hybrid ARIMA and Neural Network Model for Short-Term Price Forecasting in Deregulated Market. IEEE Transactions on Power Systems 2010; 25(1): 524-530, https://doi.org/10.1109/TPWRS.2009.2036488.

7. Bacchetti A, Saccani N. Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice. Omega 2012; 40(6): 722-737, https://doi.org/10.1016/j.omega.2011.06.008.
8. Benis A, Notea A, Barkan R. Risk and Disaster Management: From Planning and Expertise to Smart, Intelligent, and Adaptive Systems. Studies in Health Technology and Informatics 2018: 286-290.
9. Bi J, Yuan H, Zhang L, Zhang J. SGW-SCN: An integrated machine learning approach for workload forecasting in geo-distributed cloud data centers. Information Sciences 2019; 481: 57-68, https://doi.org/10.1016/j.ins.2018.12.027.
10. Bounou O, El Barkany A, El Biyaali A. Parametric Approaches for Spare Parts Demand. International Journal of Supply Chain Management 2018; 7(4): 432-439.
11. Bożejko W, Burduk A, Pempera J, Wodecki M. Optimization of production process for resource utilization. Archives of Civil and Mechanical Engineering 2019; 19(4): 1251-1258, https://doi.org/10.1016/j.acme.2019.07.002.
12. Burduk A, Jagodziński M. Assessment of Production System Stability with the Use of the FMEA Analysis and Simulation Models. In Jackowski K, Burduk R, Walkowiak K et al. (eds): Intelligent Data Engineering and Automated Learning - IDEAL 2015, Cham, Springer International Publishing: 2015; 9375: 216-223, https://doi.org/10.1007/978-3-319-24834-9_26.
13. Burduk A, Musiał K, Kochańska J et al. Tabu search and genetic algorithm for production process scheduling problem. Logforum 2019; 15(2): 181-189, https://doi.org/10.17270/J.LOG.2019.315.
14. Caggiano A. Manufacturing System. In The International Academy for Production Engineering, Laperrière L, Reinhart G (eds): CIRP Encyclopedia of Production Engineering, Berlin, Heidelberg, Springer Berlin Heidelberg: 2014: 830-836, https://doi.org/10.1007/978-3-642-20617-7_6562.
15. Chen T-Y, Lin W-T, Sheu C. A Dynamic Failure Rate Forecasting Model for Service Parts Inventory. Sustainability 2018; 10(7): 2408, https://doi.org/10.3390/su10072408.
16. Chlebus E, Helman J, Olejarczyk M, Rosienkiewicz M. A new approach on implementing TPM in a mine - A case study. Archives of Civil and Mechanical Engineering 2015; 15(4): 873-884, https://doi.org/10.1016/j.acme.2015.07.002.
17. Coleman C, Damodaran S, Chandramouli M, Deuel E. Making maintenance smarter. Predictive maintenance and the digital supply network. Deloitte University Press 2017.
18. Croston JD. Forecasting and Stock Control for Intermittent Demands. Journal of the Operational Research Society 1972; 23(3): 289-303, https://doi.org/10.1057/jors.1972.50.
19. De Gooijer JG, Hyndman RJ. 25 years of time series forecasting. International Journal of Forecasting 2006; 22(3): 443-473, https://doi.org/10.1016/j.ijforecast.2006.01.001.
20. De Livera AM, Hyndman RJ, Snyder RD. Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing. Journal of the American Statistical Association 2011; 106(496): 1513-1527, https://doi.org/10.1198/jasa.2011.tm09771.
21. Dudek-Dyduch E, Tadeusiewicz R, Horzyk A. Neural network adaptation process effectiveness dependent of constant training data availability. Neurocomputing 2009; 72(13): 3138-3149, https://doi.org/10.1016/j.neucom.2009.03.017.
22. Frank A G, Dalenogare L S, Ayala NF. Industry 4.0 technologies: Implementation patterns in manufacturing companies. International Journal of Production Economics 2019; 210: 15-26, https://doi.org/10.1016/j.ijpe.2019.01.004.
23. Ghobbar AA, Friend CH. Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model. Computers & Operations Research 2003; 30(14): 2097-2114, https://doi.org/10.1016/S0305-0548(02)00125-9.
24. Gopalakrishnan P, Banerji AK. Maintenance and spare parts management. 8. printing. New Delhi, PHI Learning - Private Limited: 2011.
25. Górnicka D, Kochańska J, Burduk A. Production Resources Utilization Improvement with the Use of Simulation Modelling. In Świątek J, Borzemski L, Wilimowska Z (eds): Information Systems Architecture and Technology: Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology - ISAT 2019, Cham, Springer International Publishing: 2020: 41-50, https://doi.org/10.1007/978-3-030-30604-5_4.
26. Gutierrez RS, Solis AO, Mukhopadhyay S. Lumpy demand forecasting using neural networks. International Journal of Production Economics 2008; 111(2): 409-420, https://doi.org/10.1016/j.ijpe.2007.01.007.
27. Hadi Hoseinie S, Ataei M, Khalokakaie R et al. Reliability analysis of drum shearer machine at mechanized longwall mines. Journal of Quality in Maintenance Engineering 2012; 18(1): 98-119, https://doi.org/10.1108/13552511211226210.
28. Hajirahimi Z, Khashei M. Hybrid structures in time series modeling and forecasting: A review. Engineering Applications of Artificial Intelligence 2019; 86: 83-106, https://doi.org/10.1016/j.engappai.2019.08.018.
29. Hall O P. Artificial Intelligence Techniques Enhance Business Forecasts. Computer-based analysis increases accuracy. The Graziadio Business Review 2002.
30. Hua ZS, Zhang B, Yang J, Tan DS. A new approach of forecasting intermittent demand for spare parts inventories in the process industries. Journal of the Operational Research Society 2007; 58(1): 52-61, https://doi.org/10.1057/palgrave.jors.2602119.
31. Hua Z, Zhang B. A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts. Applied Mathematics and Computation 2006; 181(2): 1035-1048, https://doi.org/10.1016/j.amc.2006.01.064.
32. Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: Theory and applications. Neurocomputing 2006; 70(1-3): 489-501, https://doi.org/10.1016/j.neucom.2005.12.126.
33. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. last accessed on 26.01.2021. 2018.
34. Jaganathan S, Prakash PKS. A combination-based forecasting method for the M4-competition. International Journal of Forecasting 2020; 36(1): 98-104, https://doi.org/10.1016/j.ijforecast.2019.03.030.
35. Kang R, Wang J, Cheng J et al. Intelligent forecasting of automatic train protection system failure rate in China high-speed railway. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(4): 567-576, https://doi.org/10.17531/ein.2019.4.5.
36. Khashei M, Bijari M, Hejazi SR. Combining seasonal ARIMA models with computational intelligence techniques for time series forecasting. Soft Computing 2012; 16(6): 1091-1105, https://doi.org/10.1007/s00500-012-0805-9.
37. Klindokmai S, Neech P, Wu Y et al. Evaluation of forecasting models for air cargo. The International Journal of Logistics Management 2014; 25(3): 635-655, https://doi.org/10.1108/IJLM-05-2013-0049.
38. Kourentzes N, Petropoulos F. Forecasting with multivariate temporal aggregation: The case of promotional modelling. International Journal of Production Economics 2016; 181: 145-153, https://doi.org/10.1016/j.ijpe.2015.09.011.

39. Kourentzes N, Trapero JR, Barrow DK. Optimising forecasting models for inventory planning. International Journal of Production Economics 2019: 107597, https://doi.org/10.1016/j.ijpe.2019.107597.
40. Kowalski A, Rosienkiewicz M. ANN-Based Hybrid Algorithm Supporting Composition Control of Casting Slip in Manufacture of Ceramic Insulators. In Graña M, López-Guede JM, Etxaniz O et al. (eds): International Joint Conference SOCO'16-CISIS'16-ICEUTE'16, Cham, Springer International Publishing: 2017; 527: 357-365, https://doi.org/10.1007/978-3-319-47364-2_34.
41. Kozłowski E, Mazurkiewicz D, Żabiński T et al. Assessment model of cutting tool condition for real-time supervision system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(4): 679-685, https://doi.org/10.17531/ein.2019.4.18.
42. Kozłowski E, Mazurkiewicz D, Żabiński T et al. Machining sensor data management for operation-level predictive model. Expert Systems with Applications 2020; 159: 113600, https://doi.org/10.1016/j.eswa.2020.113600.
43. Kozłowski T, Wodecki J, Zimroz R et al. A Diagnostics of Conveyor Belt Splices. Applied Sciences 2020; 10(18): 6259, https://doi.org/10.3390/app10186259.
44. Li R, Jiang P, Yang H, Li C. A novel hybrid forecasting scheme for electricity demand time series. Sustainable Cities and Society 2020; 55: 102036, https://doi.org/10.1016/j.scs.2020.102036.
45. Li Y, Wang K. Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 22(1): 63-72, https://doi.org/10.17531/ein.2020.1.8.
46. Liang T-F. Application of fuzzy sets to manufacturing/distribution planning decisions in supply chains. Information Sciences 2011; 181(4): 842-854, https://doi.org/10.1016/j.ins.2010.10.019.
47. Lucas Silva A, Ribeiro R, Teixeira M. Modeling and control of flexible context-dependent manufacturing systems. Information Sciences 2017; 421: 1-14, https://doi.org/10.1016/j.ins.2017.08.084.
48. Makridakis S, Spiliotis E, Assimakopoulos V. The M4 Competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting 2020; 36(1): 54-74, https://doi.org/10.1016/j.ijforecast.2019.04.014.
49. Manafzadeh Dizbin N, Tan B. Optimal control of production-inventory systems with correlated demand inter-arrival and processing times. International Journal of Production Economics 2020; 228: 107692, https://doi.org/10.1016/j.ijpe.2020.107692.
50. Manzke L, Keller B, Buscher U. An Artificial Bee Colony Algorithm to Solve the Single Row Layout Problem with Clearances. In Wilimowska Z, Borzemski L, Świątek J (eds): Information Systems Architecture and Technology: Proceedings of 38th International Conference on Information Systems Architecture and Technology - ISAT 2017, Cham, Springer International Publishing: 2018; 657: 285-294, https://doi.org/10.1007/978-3-319-67223-6_27.
51. Mulay A, Ben BS, Ismail S, Kocanda A. Prediction of average surface roughness and formability in single point incremental forming using artificial neural network. Archives of Civil and Mechanical Engineering 2019; 19(4): 1135-1149, https://doi.org/10.1016/j.acme.2019.06.004.
52. Omar H, Hoang VH, Liu D-R. A Hybrid Neural Network Model for Sales Forecasting Based on ARIMA and Search Popularity of Article Titles. Computational Intelligence and Neuroscience 2016; 2016: 1-9, https://doi.org/10.1155/2016/9656453.
53. Ömer Faruk D. A hybrid neural network and ARIMA model for water quality time series prediction. Engineering Applications of Artificial Intelligence 2010; 23(4): 586-594, https://doi.org/10.1016/j.engappai.2009.09.015.
54. Ou D, Tang M, Xue R, Yao H. Hybrid fault diagnosis of railway switches based on the segmentation of monitoring curves. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 514-522, https://doi.org/10.17531/ein.2018.4.2.
55. Pereira DF, Oliveira JF, Carravilla MA. Tactical sales and operations planning: A holistic framework and a literature review of decision-making models. International Journal of Production Economics 2020; 228: 107695, https://doi.org/10.1016/j.ijpe.2020.107695.
56. Pérez-Chacón R, Asencio-Cortés G, Martínez-Álvarez F, Troncoso A. Big data time series forecasting based on pattern sequence similarity and its application to the electricity demand. Information Sciences 2020; 540: 160-174, https://doi.org/10.1016/j.ins.2020.06.014.
57. Petropoulos F, Kourentzes N, Nikolopoulos K. Another look at estimators for intermittent demand. International Journal of Production Economics 2016; 181: 154-161, https://doi.org/10.1016/j.ijpe.2016.04.017.
58. Pham HT, Tran VT, Yang B-S. A hybrid of nonlinear autoregressive model with exogenous input and autoregressive moving average model for long-term machine state forecasting. Expert Systems with Applications 2010; 37(4): 3310-3317, https://doi.org/10.1016/j.eswa.2009.10.020.
59. Praekhaow P. Determination of Trading Points using the Moving Average Methods. Bangkok, Thailand, 2010; GMSTEC: 6.
60. Rathod S, Mishra GC, Singh KN. Hybrid Time Series Models for Forecasting Banana Production in Karnataka State, India. Journal of the Indian Society of Agricultural Statistics 2017: 9.
61. Rego JR do, Mesquita MA de. Spare parts inventory control: a literature review. Production 2011; 21(4): 645-666.
62. Rojek I, Kowal M, Stoic A. Predictive compensation of thermal deformations of ball screws in CNC machines using neural networks. Tehnicki vjesnik - Technical Gazette 2017, https://doi.org/10.17559/TV-20161207171012.
63. Romański L, Bieniek J, Komarnicki P et al. Estimation of operational parameters of the counter-rotating wind turbine with artificial neural networks. Archives of Civil and Mechanical Engineering 2017; 17(4): 1019-1028, https://doi.org/10.1016/j.acme.2017.04.010.
64. Rosienkiewicz M. Accuracy Assessment of Artificial Intelligence-Based Hybrid Models for Spare Parts Demand Forecasting in Mining Industry. In Wilimowska Z, Borzemski L, Świątek J (eds): Information Systems Architecture and Technology: Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology - ISAT 2019, Cham, Springer International Publishing: 2020; 1052: 176-187, https://doi.org/10.1007/978-3-030-30443-0_16.
65. Rosienkiewicz M, Chlebus E, Detyna J. A hybrid spares demand forecasting method dedicated to mining industry. Applied Mathematical Modelling 2017; 49: 87-107, https://doi.org/10.1016/j.apm.2017.04.027.
66. Rosienkiewicz M, Kowalski A, Helman J, Zbieć M. Development of Lean Hybrid Furniture Production Control System based on Glenday Sieve, Artificial Neural Networks and Simulation Modeling. Drvna industrija 2018; 69(2): 163-173, https://doi.org/10.5552/drind.2018.1747
67. Ruiz-Aguilar JJ, Turias IJ, Jiménez-Come MJ. Hybrid approaches based on SARIMA and artificial neural networks for inspection time series forecasting. Transportation Research Part E: Logistics and Transportation Review 2014; 67: 1-13, https://doi.org/10.1016/j.tre.2014.03.009.
68. Schwarz G. Estimating the Dimension of a Model. The Annals of Statistics 1978; 6(2): 461-464, https://doi.org/10.1214/aos/1176344136.
69. Segovia Ramirez I, Mohammadi-Ivatloo B, García Márquez FP. Alarms management by supervisory control and data acquisition system for wind turbines. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(1): 110-116, https://doi.org/10.17531/ein.2021.1.12.

70. Sekala A, Gwiazda A, Kost G, Banas W. Modelling of a production system using the multi-agent network approach. IOP Conference Series: Materials Science and Engineering 2018; 400: 052009, https://doi.org/10.1088/1757-899X/400/5/052009.
71. Seliger G. Maintenance. In The International Academy for Production Engineering, Laperrière L, Reinhart G (eds): CIRP Encyclopedia of Production Engineering, Berlin, Heidelberg, Springer Berlin Heidelberg: 2014: 818-821, https://doi.org/10.1007/978-3-642-20617-7_12.
72. Sikder S, Mukherjee I, Panja SC. A synergistic Mahalanobis-Taguchi system and support vector regression based predictive multivariate manufacturing process quality control approach. Journal of Manufacturing Systems 2020; 57: 323-337, https://doi.org/10.1016/j.jmsy.2020.10.003.
73. Sinisterra WQ, Cavalcante CAV. An integrated model of production scheduling and inspection planning for resumable jobs. International Journal of Production Economics 2020; 227: 107668, https://doi.org/10.1016/j.ijpe.2020.107668.
74. Smyl S. A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. International Journal of Forecasting 2020; 36(1): 75-85, https://doi.org/10.1016/j.ijforecast.2019.03.017.
75. Sobaszek Ł, Gola A, Świć A. Time-based machine failure prediction in multi-machine manufacturing systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 22(1): 52-62, https://doi.org/10.17531/ein.2020.1.7.
76. Suomala P, Sievanen M, Paranko J. The effects of customization on spare part business: A case study in the metal industry. International Journal of Production Economics 2002; 79(1): 57-66, https://doi.org/10.1016/S0925-5273(00)00060-8.
77. Valis D, Forbelská M, Vintr Z. Forecasting study of mains reliability based on sparse field data and perspective state space models. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(2): 179-191, https://doi.org/10.17531/ein.2020.2.1.
78. Van der Auweraer S, Boute R. Forecasting spare part demand using service maintenance information. International Journal of Production Economics 2019; 213: 138-149, https://doi.org/10.1016/j.ijpe.2019.03.015.
79. Van Horenbeek A, Buré J, Cattrysse D et al. Joint maintenance and inventory optimization systems: A review. International Journal of Production Economics 2013; 143(2): 499-508, https://doi.org/10.1016/j.ijpe.2012.04.001.
80. Wagner SM, Jönke R, Eisingerich AB. A Strategic Framework for Spare Parts Logistics. California Management Review 2012; 54(4): 69-92, https://doi.org/10.1525/cmr.2012.54.4.69.
81. Wan C, Xu Z, Wang Y et al. A Hybrid Approach for Probabilistic Forecasting of Electricity Price. IEEE Transactions on Smart Grid 2014; 5(1): 463-470, https://doi.org/10.1109/TSG.2013.2274465.
82. Więcek D, Burduk A, Kuric I. The use of ANN in improving efficiency and ensuring the stability of the copper ore mining process. Acta Montanistica Slovaca 2019; 24(1): 14.
83. Yang L, Li B. The Combination Forecasting Model of Grain Production Based on Stepwise Regression Method and RBF Neural Network. Advance Journal of Food Science and Technology 2015; 7(11): 891-895, https://doi.org/10.19026/ajfst.7.2528.
84. Yin S, Liu L, Hou J. A multivariate statistical combination forecasting method for product quality evaluation. Information Sciences 2016; 355-356: 229-236, https://doi.org/10.1016/j.ins.2016.03.035.
85. Yu L, Liang S, Chen R, Lai KK. Predicting monthly biofuel production using a hybrid ensemble forecasting methodology. International Journal of Forecasting 2019. doi:10.1016/j.ijforecast.2019.08.014, https://doi.org/10.1016/j.ijforecast.2019.08.014.
86. Zhang X, Wang J, Gao Y. A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM. Energy Economics 2019; 81: 899-913, https://doi.org/10.1016/j.eneco.2019.05.026.
87. Zhou J, Qiu Y, Zhu S et al. Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate. Engineering Applications of Artificial Intelligence 2021; 97: 104015, https://doi.org/10.1016/j.engappai.2020.104015.
88. https://www.pwc.pl/pl/pdf/industry-4-0.pdf, last accessed 26.01.2021.
89. http://www.mesasoftware.com/papers/ZeroLag.pdf, last accessed 10.02.2020.
90. http://documentation.statsoft.com/STATISTICAHelp.aspx?path=MachineLearning/MachineLearning/Overviews/, SupportVectorMachinesIntroductoryOverview,last accessed 18.02.2020.

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl
Article citation info:
Sun B, Yang X, Ren Y, Wang Z, Antosz K, Loska A, Jasiulewicz-Kaczmarek M. Failure-based sealing reliability analysis considering dynamic interval and hybrid uncertainties. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 278–284, http://doi.org/10.17531/ein.2021.2.7.

Indexed by: Web of Science Group

Failure-based sealing reliability analysis considering dynamic interval and hybrid uncertainties

Bo Sun [a], Xi Yang [a], Yi Ren [a], Zili Wang [a], Katarzyna Antosz [b], Andrzej Loska [c], Małgorzata Jasiulewicz-Kaczmarek [d]

[a] School of Reliability and Systems Engineering, Beihang University (BUAA), No.37 XueYuan RD. Haidian, Beijing 100191, China
[b] Faculty of Mechanical Engineering and Aeronautics, Rzeszow University of Technology, Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
[c] Faculty of Organization and Management, Silesian University of Technology, ul. Roosevelta 26-28, 41-800 Zabrze, Poland
[d] Faculty of Management Engineering, Poznan University of Technology, ul. Prof. Rychlewskiego 2, 60-965 Poznan, Poland

Highlights

- The degradation with a small sample is modelled using a dynamic interval process.
- The reliability model with hybrid uncertainty is proposed.
- The case study of sealing reliability prediction and evaluation is presented.

Abstract

In the reliability analysis of a sealing structure, radial clearance of the contact surface is usually regarded as a failure criterion, and the sample size is usually quite small, which brings great challenges to uncertainty quantification. Therefore, this paper proposes a reliability analysis method based on the leakage mechanism of the sealing. With the application of dynamic interval, the proposed method can be used to deal with problem of degradation in small sample to evaluate reliability. Moreover, the dynamic reliability with the mixture of the probabilistic and non-probabilistic variables can be obtained using the proposed method. An illustrative numerical case study of a spool valve is conducted in order to validate the proposed method and the implemented reliability sensitivity analysis. The proposed method is of great help in evaluating and predicting reliability with small degradation sample and hybrid uncertainties.

Keywords

reliability analysis, dynamic interval, hybrid uncertainty, physics of failure, sealing.

1. Introduction

The problem of machines and devices assessment is considered as one of the most important and relevant reliability analysis issues [12, 13, 27, 29]. It is directly related to many aspects of technical systems exploitation, including efficiency and sustainability dimensions [16, 18, 19, 22]. The exploitation assessment is also a key component of an operational decision-making process as a result of the established maintenance policy [28, 37]. One of the most important features underlying the construction of the exploitation assessment models is reliability. In recent years, more and more engineers and statisticians have acquired and processed degradation data through the measurement of performance parameters of several products in order to predict their reliability. Mathematical and other solutions built on the basis of the reliability methodology are still up-to-date [22, 41], especially when it is necessary to take into account a particular degree of uncertainty [22, 38, 46].

The key element to increase the reliability and performance of mechanical devices is the structural reliability analysis. The contemporary complexity of machines and devices still makes it a big challenge for both scientists and practitioners. Because of a very large number of calculations required for assessing small failure probabilities, this is a labor-intensive and time-consuming process [30]. Design and operational parameters of mechanical elements of devices due to the effects of environmental changes are often uncertain. According to [10], different typologies of uncertainty and analysis, methods for reliability can be divided into two main categories: time-variant and time invariant methods. Therefore, the analysis of machines and devices reliability is focused on the identification and evaluation of various types of uncertainty, their effects and the assessment of the probability of a component failure [38].

These research challenges appear when spool valves are taken into consideration as an example of a mechanical element. Spool valve is a basic part in a hydraulic system, where reliability has a significant influence on the entire system [35]. The reciprocating sliding operations result in the inevitable wearing of a spool and sleeve, which leads to leakage in the sealing and eventually causes the failure of the sealing [25]. The wear degradation was investigated by Liu et al [26] and Yang et al [43]. In these studies, the failure is caused by the wear volume exceeding the threshold. Various degradation models have attracted the attention of researchers all over the world [33]. Gorjian et al [15] and Shahraki et al [32] reviewed various degradation models in a reliability analysis. Moreover, in the work [8], a probabilistic method based on a stochastic differential calculation for the reliability

E-mail addresses: B. Sun - sunbo@buaa.edu.cn, X. Yang - bryantyx@buaa.edu.cn, Y. Ren - renyi@buaa.edu.cn, Z. Wang - wzl@buaa.edu.cn, K. Antosz - katarzyna.antosz@prz.edu.pl, A. Loska - andrzej.loska@polsl.pl, M. Jasiulewicz-Kaczmarek - malgorzata.jasiulewicz-kaczmarek@put.poznan.pl

assessment of structural components is defined. Andrieu-Renaud et al [1] developed a method known as PHI2, based on a cross approach that solves reliability problems using classic time-invariant reliability tools. In the aspect of a modern performance-based design, Au and Beck [2] implemented Subset Simulation (SS) to evaluate the performance of structures. The SS method, for the assessment of small failure probabilities, was also used by Bourinet et al [5] in the approach referred as 2SMART.

As a matter of fact, the failure of the sealing of a spool valve is caused by internal leakage [31], whose mechanism shall be taken into account in order to conduct the reliability analysis [11]. The leakage between contacting surfaces in valves is likely to be influenced by the direction of the surface anisotropy or lay [3, 23]. A quantitative multi-scale analysis of the surface morphology or curvature are considered as valuable tools for elucidating changes in the anisotropy caused by processing, and for the performance indication including sealing, lubrication, and friction [4].

The existence of degradation brings the issue of time-variant reliability [25], which indicates that the reliability of the seal varies with a task. In practice, the wear data is usually hard to measure [1] and the observed data is often insufficient to quantify the uncertainty using a probabilistic approach [42]. Therefore, the interval method is applied to deal with the problem of a small sample [20]. In order to solve this problem, Liao et al [24] built a reliability model of aviation seal with an interval method. The interval method is able to quantify the uncertainty with limited data, and has been eagerly wildly applied in the field of epistemic uncertainty and non-probabilistic reliability. Kang et al [21] proposed belief reliability as a reliability metric under the epistemic uncertainty. You et al [45] presented a novel structural reliability analysis method with fuzzy random variables. Besides, more than one type of uncertainties exist due to the interval and random variables [36]. Hence, the reliability modelling with the hybrid uncertainty has become another research focus in recent years [6, 17]. Moreover, Wang [39], according to [40, 47], defined a hybrid reliability analysis (HRA) as a task that quantifies two types of uncertainties, and as a core one in the structural reliability research. Chakraborty et al [7] analyzed structural probabilistic safety under the hybrid uncertainty. Sun et al [34] built the time-variant reliability model using the hybrid non-probability method. Jiang et al [20] reviewed several main research directions in the probability-interval hybrid uncertainty analysis, and provided an outlook for the potential research aspects. A Bayesian approach for a sealing failure analysis was presented in [44], where the radial clearance height was regarded as a failure criterion, and where the observed data was assumed to follow the gamma distribution.

However, in the reliability analysis of the sealing, a performance function shall be established based on a leakage mechanism. The sample of degradation data is usually very small, and the existing research cannot solve the problem of a small sample with degradation. Therefore, this paper proposed a physics of a failure-based reliability analysis model, where the dynamic interval is applied to deal with the issue of degradation in a small sample. The proposed method takes a failure mechanism into consideration to build a dynamic reliability model, and the reliability is resolved with the hybrid uncertainty method. Moreover, an illustrative case study of the sealing in a spool valve is conducted to validate the proposed method and to analyse the dynamic reliability of the sealing.

2. Dynamic reliability model considering degradation and hybrid uncertainty

For the need of the dynamic reliability model design, it's necessary to introduce an uncertainty process and, next, a dynamic interval process. The uncertainty process can be expressed as:

$$\left\{X\left(t_i\right)\in X^I\left(t\right),t\in T\right\} \tag{1}$$

where: $X\left(t_i\right)$ denotes an interval value at a given time t_i. For the time $t_1,t_2,\ldots,t_n$, the joint distribution region composed of interval variables is a hypercube domain. Therefore, the uncertainty process is defined as a dynamic interval process [38].

The dynamic interval can be described as a time-variant interval process where an interval changes with time. For the given interval process $X\left(t\right)$, $\overline{X\left(t\right)}$ and $\underline{X\left(t\right)}$ denote the upper and lower limits, respectively, and the mean function of the interval process is expressed as:

$$X^c=\frac{\overline{X\left(t\right)}+\underline{X\left(t\right)}}{2} \tag{2}$$

and the radius function is expressed as:

$$X^r=\frac{\overline{X\left(t\right)}-\underline{X\left(t\right)}}{2} \tag{3}$$

Once X^r and X^c are obtained, the uncertainty characteristics of each specific moment can be determined. The mean and radius functions can be obtained using the fitting methods, such as a linear model, an exponential model or a stochastic process model. The dynamic interval process requires the interval information at each observation point, which enables the proposed method to appraise the dynamic reliability with very limited data. Furthermore, the fitted curve with the observation can be used to predict the reliability in a longer period.

When both the random uncertainty and interval uncertainty exist at the same time in a structure, the performance function can be expressed as:

$$Z=g\left(X,Y\right) \tag{4}$$

where: $g\left(\cdot\right)$ denotes the function of X and Y,
$X=\left\{X_1,X_2,\ldots,X_n\right\}$ denotes independent *n*-dimensional interval vector,
$X_i\in X_i^1=\left[X_i^L,X_i^U\right],i=1,2,\ldots,n$,
$Y=\left\{Y_1,Y_2,\ldots Y_m\right\}$ denotes independent *m*-dimensional random vector.

The structure is considered as reliable with Z>0, and unreliable with Z<0. In practice, Z is the difference between the performance threshold and performance parameters, which can be expressed as:

$$Z=P_{th}-P\left(x_1,x_2,\ldots,x_n\right) \tag{5}$$

where: $P\left(x_1,x_2,\ldots,x_n\right)$ is the function of critical performance parameters modelled with a failure mechanism of a product,
P_{th} denotes the performance threshold.

The reliability is expressed as:

$$R=P\left\{g\left(X,Y\right)>0\right\} \tag{6}$$

The reliability result becomes an interval instead of a probability value due to the existence of interval vectors, of which the lower limit R_L and upper limit R_U are given by the following equations:

$$R_L = P\{\min g(X,Y) > 0\}$$

$$R_U = P\{\max g(X,Y) > 0\} \tag{7}$$

When interval vectors or random vectors are time-dependent, the reliability given in (7) becomes a time-variant interval with $R_L(t)$ and $R_U(t)$. In the proposed model, the leakage mechanism of the sealing in a spool valve with necessary features is presented to determine the reliability.

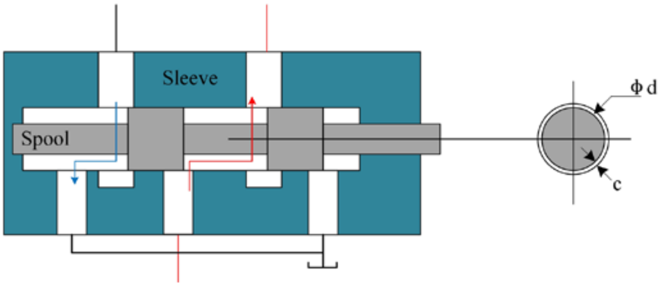


Fig. 1. Schema of a spool valve and clearance [44]

A typical directional valve is shown in Fig. 1. As discussed in [44], an internal leakage is given by the following equation:

$$Q = \Delta P \frac{\pi d c^3}{12 \mu L} \tag{8}$$

where: Q - internal leakage of a spool valve,

ΔP - pressure difference,

d - diameter of a spool valve,

c - radial clearance height,

μ - dynamic viscosity of hydraulic oil,

L - clearance length.

The high frequency of the back and forth sliding movements cause the wear of spools and sleeves [14]. The wear will finally cause the increase of the clearance and internal leakage exceeding its allowable threshold. Thus, the sealing of the spool valve is regarded as a failure. Therefore, according to formula (5), the performance function of the sealing is given by:

$$Z = Q_{th} - \Delta P \frac{\pi d c^3}{12 \mu L} \tag{9}$$

where: Q_{th} - the threshold of leakage.

For each observation, the maximum and minimum are regarded as interval limits:

$$c^I_{n_i} = [c^L, c^U] = [\min(S(i)), \max(S(i))] \tag{10}$$

where: $S(i)$ - the samples from the i_{th} observation,

c^L, c^U - interval limits.

Based on the performance function in (9), the reliability can be calculated using the first order second moment (FOSM) method. In this circumstance, Z is the function of $Q_{th}, \Delta P, d, \mu, L, c$. Z varies with the stroke, which is expressed as:

$$Z(n) = g(Q_{th}, \Delta P, d, \mu, L, c(n)) \tag{11}$$

In formula (11), $c(n)$ varies with the stroke and is modelled with a dynamic interval process. The reliability decreases monotonically with the increase of c. Therefore, the dynamic reliability interval of the sealing can be obtained using the following formulas:

$$R^L(n) = P\{g(Q_{th}, \Delta P, d, \mu, L, c) > 0 \mid c = c^U(n)\}$$

$$R^U(n) = P\{g(Q_{th}, \Delta P, d, \mu, L, c) > 0 \mid c = c^L(n)\} \tag{12}$$

3. Numerical example

In the case study, the variables $Q_{th}, \Delta P, d, \mu, L$ (9) are regarded as random variables, and the distributions of the random variables are assumed to be normal, as shown in Table 1. The clearance of the spool and sleeve c is usually measurable with a small sample which is regarded as an interval variable and described with a dynamic interval process.

For the purposes of determining the dynamic interval of clearance, the wear volumes are observed per 50,000 strokes. The clearance of a

Table 1. Random variables in sealing

Parameter	Type of distribution	Mean	Coefficient of variation
Q_{th}	Normal distribution	14.54 ml/min	0.03
ΔP	Normal distribution	27.50 MPa	0.03
d	Normal distribution	7.60 mm	0.03
μ	Normal distribution	0.013 kg/(m·s)	0.03
L	Normal distribution	0.50 mm	0.03

Table 2. Wear volumes and interval of clearance[44]

Unit	**Number of Strokes (10 thousands)**																			
	5	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**	**100**
Sample1	9	13	25	28	31	35	41	45	49	56	61	63	69	76	81	85	92	96	104	113
Sample2	5	10	17	24	30	39	44	51	58	63	68	76	80	85	93	99	104	108	113	119
Sample3	6	12	18	23	26	30	38	43	49	57	63	70	74	77	84	92	99	101	107	113
Sample4	3	6	17	20	25	33	41	46	54	60	64	69	72	83	89	93	96	100	104	111
Sample5	10	15	19	24	34	43	49	54	61	65	70	74	77	82	84	91	99	103	108	115
c^U	10	15	25	28	34	43	49	54	61	65	70	76	80	85	93	99	104	108	113	119
c^L	3	6	17	20	25	30	38	43	49	56	61	63	69	76	81	85	92	96	104	111

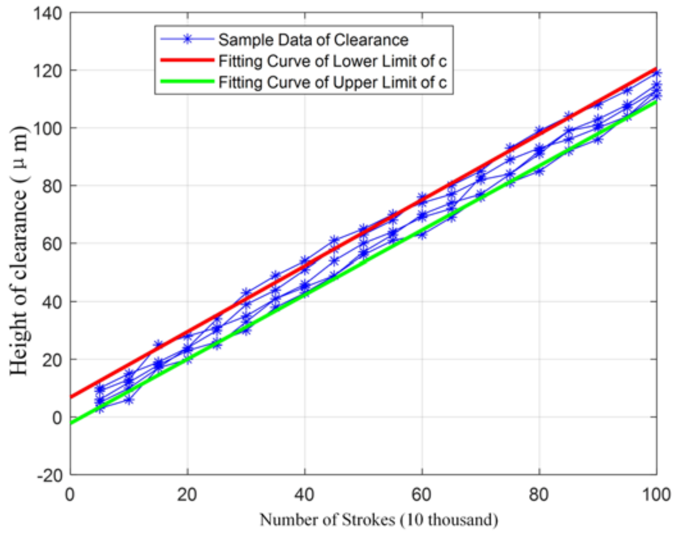


Fig. 2. Sample data of clearance and fitting limits

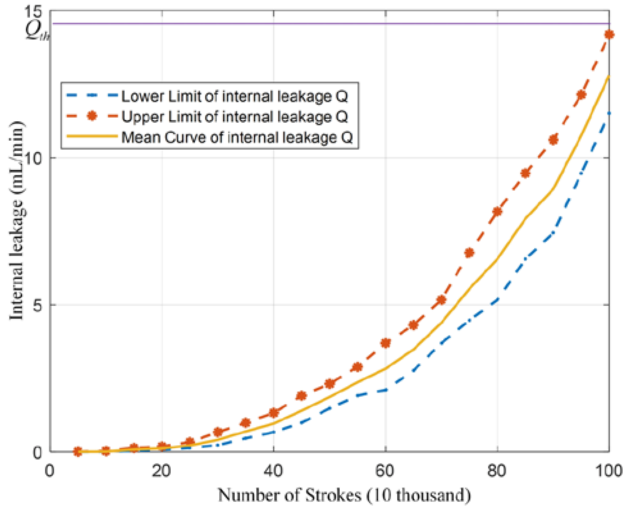


Fig. 3. The trend of internal leakage

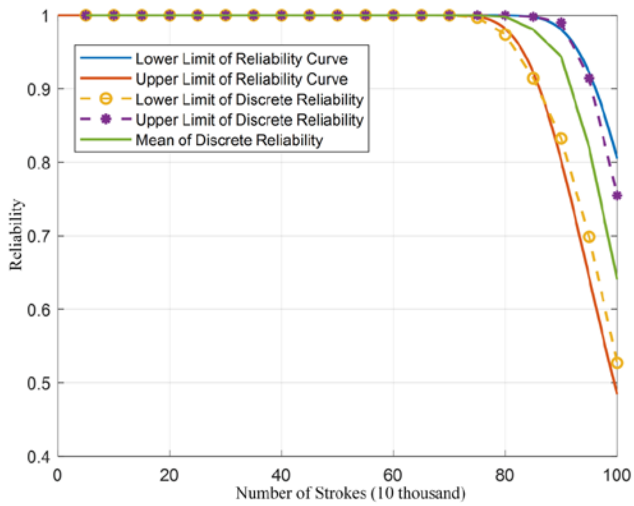


Fig. 4. Dynamic reliability of the sealing varies with the stroke

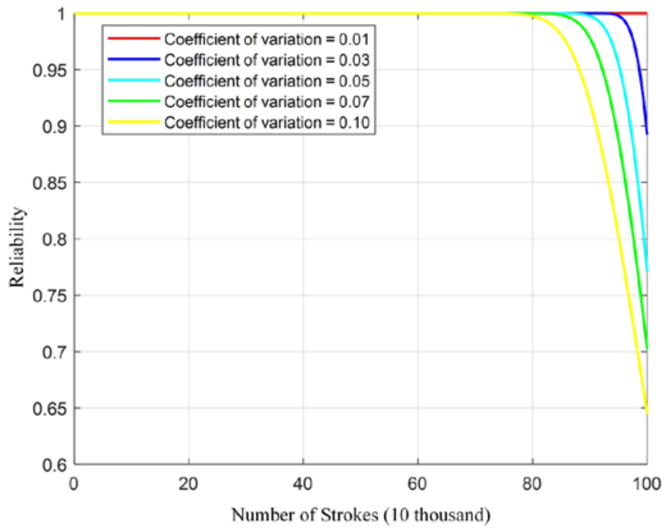


Fig. 5. Dynamic reliability with various coefficients of the variation

spool valve of 5 samples is listed in Table 2, followed by the interval limits of c^L and c^U (10).

With the interval data of clearance, the trend of the upper and lower limits can be obtained using the curve fitting method, such as the least square method. The functions of c^L and c^U vary with the strokes n and are fitting curves as:

$$\begin{aligned} c^L(n) &= 1.113n - 2.179 \\ c^U(n) &= 1.138n + 6.811 \end{aligned} \tag{12}$$

where n denotes the number of strokes with a unit of 10 thousand. The initial sample data and the fitting curve of limits are depicted in Fig. 2.

With the increase of clearance, the internal leakage also ascends with the strokes. The leakage also becomes an interval variable due to the interval uncertainty of clearance. The limits and mean curves are depicted in Fig. 3. With the stroke of the spool valve increase, the internal leakage gets closer to a failure criterion, which will cause the descending of the sealing reliability.

The sealing reliability from the discrete interval and fitting dynamic interval of c are depicted according to (12) in Fig. 4.

It can be concluded from the performance function that $\Delta P, d$ and c have a negative effect on the reliability of the sealing, while Q_{th}, μ and L have a positive effect. In addition, the variation of all the parameters in the spool valve will also influence the reliability trend of the sealing.

In order to compare the differences caused by the variation of the parameters, dynamic curves with various coefficients of the variation are depicted in Fig. 5. It can be seen from the curves in Fig. 5 that with the increase of the variation, the reliability of the sealing will decrease earlier with a lower terminal point. Therefore, it is suggested that the inconsistency and uncertainty shall be reduced to obtain a longer lifetime with higher reliability. It can be seen from the curves in Fig. 5 that with the increase of the variation, the reliability of the sealing will decrease earlier with a lower terminal point. Therefore, it is suggested that the inconsistency and uncertainty shall be reduced to obtain a longer lifetime with higher reliability.

As it was mentioned above, the proposed method enables the reliability prediction with less data. Therefore, the first 15 observed data in Table 2 was used to build a dynamic reliability model with the proposed method. Additionally, the reliability curves with less observations and full observations are depicted in Fig. 6.

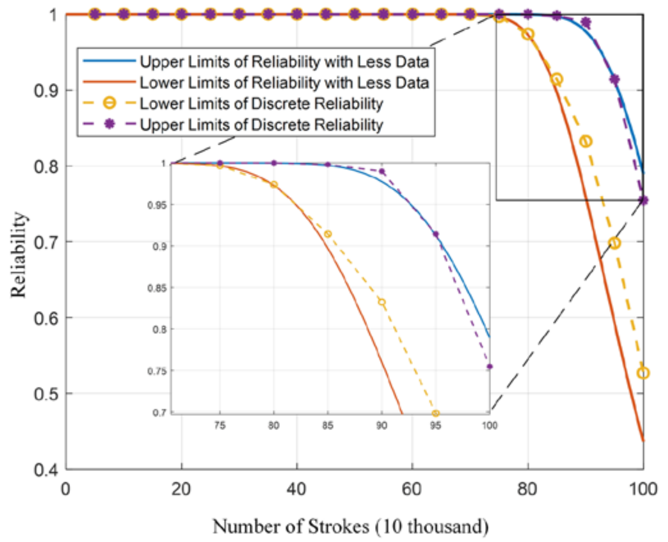


Fig. 6. Comparison between the predicted reliability curves and the evaluated reliability value

Table 3. The predicted and evaluated reliability

Items	Number of Strokes (10 thousand)				
	80	85	90	95	100
Predicted upper limits	0.9998	0.9970	0.9774	0.9125	0.7894
Evaluated upper limits	0.9999	0.9980	0.9899	0.9143	0.7548
Relative Error	0.01%	0.10%	1.2%	0.19%	4.5%
Predicted lower limits	0.9726	0.8968	0.7598	0.5931	0.4366
Evaluated lower limits	0.9739	0.9143	0.8323	0.6984	0.5270
Relative Error	0.13%	1.9%	8.7%	15%	17%

As shown in Fig. 6, the predicted reliability is basically consistent with the evaluated reliability., The numerical comparison results are listed in Table 3.

The last 5 groups of the data (observed at 80 to 100) in Table 2 are used to verify the proposed method. The relative error at each observation point is very small. It can be seen from the comparison in Table 3 that the predicted reliability limits with partial data are quite close to the evaluated limits with full data, revealing that the proposed method is very efficient in predicting the reliability limits with the insufficient observed data.

4. Conclusions

This paper proposed a failure-based reliability analysis method for sealing. In the reliability analysis of sealing, the allowable leakage is regarded as a failure criterion. The proposed method takes a failure mechanism of sealing into consideration and establishes a performance function of sealing with an explicit expression with which the sensitivity of different variables can be easily obtained. Besides, a dynamic interval is adapted to deal with the issues of a small sample degradation. The reliability can be evaluated with the hybrid uncertainty method. The obtained reliability result becomes two boundary curves instead of one reliability curve due to the existence of both interval variables and random variables.

The proposed method can be used to predict the wear interval in sealing with small sample and hybrid uncertainties, and the reliability can be evaluated and predicted with limited observation data. The operators can make dependable maintenance decision with the reliability trend curve, thus replacing the sealing when reliability drops to a certain level.

Moreover, an illustrative case study is conducted to verify the proposed method, where the data in [44] is applied to build a dynamic reliability model. The method proposed in this paper is more concise and confident under the circumstance of only a small amount of data than that with the gamma process and Bayesian estimation. Furthermore, it is verified if the proposed method can be used to predict the reliability with high precision in case of partial data, as well as to evaluate the reliability with a reasonable interval result in case of full data. The following can be concluded from the previous discussions:

1. The proposed method can deal with the problem of degradation with a small sample, which is common in many engineering practices. The dynamic interval process can be applied to quantify the dynamic uncertainty with insufficient data.
2. A failure mechanism shall be taken into account when establishing the performance function. The influences and sensitivity of different parameters can be obtained easily with an explicit expression.
3. The hybrid uncertainty problem is transformed into the probabilistic reliability in the proposed method, with which the reliability boundary curves can be obtained.
4. Compared with the approach of the reliability analysis in [44], the proposed method has the advantage of simplicity and credibility without any subjective hypothesis for the observed data.

References

1. Andrieu-Renaud C, Sudret B, Lemaire M. The PHI2 method: a way to compute time-variant reliability. Reliability Engineering & System Safety 2004; 84(1): 75-86, https://doi.org/10.1016/j.ress.2003.10.005.
2. Au SK, Beck JL. Subset simulation and its application to seismic risk based on dynamic analysis. Journal of Engineering Mechanics 2003; 129(8): 901-917, https://doi.org/10.1061/(ASCE)0733-9399(2003)129:8(901).
3. Bartkowiak T, Berglund J, Brown CA. Establishing functional correlations between multiscale areal curvatures and coefficients of friction for machined surfaces. Surface Topography: Metrology and Properties 2018; 6(3): 034002, https://doi.org/10.1088/2051-672X/aac073.
4. Bartkowiak T, Berglund J, Brown CA. Multiscale characterizations of surface. Materials 2020; 13(13): 3028, https://doi.org/10.3390/ma13133028.
5. Bourinet JM, Deheeger F, Lemaire M. Assessing small failure probabilities by combined subset simulation and support vector machines. Structural Safety 2011; 33(6): 343-353, https://doi.org/10.1016/j.strusafe.2011.06.001.
6. Brunskill H, Harper P, Lewis R. The real-time measurement of wear using ultrasonic reflectometry. Wear 2015; 332-333: 1129-1133, https://doi.org/10.1016/j.wear.2015.02.049.
7. Chakraborty S, Sam P C. Probabilistic safety analysis of structures under hybrid uncertainty. International Journal for Numerical Methods in Engineering 2007; 70(4): 405-422, https://doi.org/10.1002/nme.1883.
8. Ciampoli M. Time dependent reliability of structural systems subject to deterioration. Computers Structures 1998; 67(1-3): 29-35, https://doi.org/10.1016/S0045-7949(97)00153-3.

9. Du X. Interval reliability analysis. International Design Engineering Technical Conferences and Computers and Information in Engineering Conference 2007; 48078: 1103-1109, https://doi.org/10.1115/DETC2007-34582.
10. En XN, Zhang YM, Huang XZ. Time-variant reliability analysis of a continuous system with strength deterioration based on subset simulation. Advances in Manufacturing 2019; 7(2): 188-198, https://doi.org/10.1007/s40436-019-00252-7.
11. Fan M, Zeng Z, Kang R, et al. Reliability modeling of a spool valve considering dependencies among failure mechanisms. Safety and Reliability of Complex Engineered Systems-Proceedings of the 25th European Safety and Reliability Conference, ESREL 2015: 4407-4411, https://doi.org/10.1201/b19094-579.
12. Franciosi C, Voisin A, Miranda S, Riemma S, Iung, B. Measuring maintenance impacts on sustainability of manufacturing industries: from a systematic literature review to a framework proposal. Journal of Cleaner Production 2020; 260: 121065, https://doi.org/10.1016/j.jclepro.2020.121065.
13. Fraser K, Hvolby HH, Tseng TL. Maintenance management models: a study of the published literature to identify empirical evidence. International Journal of Quality & Reliability Management 2015; 32(6): 635-664, https://doi.org/10.1108/IJQRM-11-2013-0185.
14. Gordić D, Babić M, Milovanović D, et al. Spool valve leakage behavior. Archives of Civil and Mechanical Engineering 2011, 11(4): 859-866, https://doi.org/10.1016/S1644-9665(12)60082-X.
15. Gorjian N, Ma L, Mittinty M, et al. A review on degradation models in reliability analysis. Engineering Asset Lifecycle Management. Springer, London 2010: 369-384, https://doi.org/10.1007/978-0-85729-320-6_42.
16. Holgado M, Macchi M, Evans S. Exploring the impacts and contributions of maintenance function for sustainable manufacturing, International Journal of Production Research 2020; 58: 7292-7310, https://doi.org/10.1080/00207543.2020.1808257.
17. Hurtado JE, Alvarez DA. The encounter of interval and probabilistic approaches to structural reliability at the design point. Computer Methods in Applied Mechanics and Engineering 2012; 225-228: 74-94, https://doi.org/10.1016/j.cma.2012.03.020.
18. Janik A, Ryszko A. Mapping the field of Industry 4.0 based on bibliometric analysis. Vision 2020: Sustainable Economic Development and Application of Innovation Management, 32nd Conference of the International-Business-Information-Management-Association (IBIMA) 2018: 6316-6330.
19. Jasiulewicz-Kaczmarek M, Gola A. Maintenance 4.0 Technologies for Sustainable Manufacturing - an Overview. IFAC PapersOnLine 2019; 52-10: 91-96, https://doi.org/10.1016/j.ifacol.2019.10.005.
20. Jiang C, Zheng J, Han X. Probability-interval hybrid uncertainty analysis for structures with both aleatory and epistemic uncertainties: a review. Structural and Multidisciplinary Optimization 2018; 57(6): 2485-2502, https://doi.org/10.1007/s00158-017-1864-4.
21. Kang R, Zhang Q, Zeng Z, et al. Measuring reliability under epistemic uncertainty: Review on non-probabilistic reliability metrics. Chinese Journal of Aeronautics 2016; 29(3): 571-579, https://doi.org/10.1016/j.cja.2016.04.004.
22. Kozłowski E, Mazurkiewicz D, Kowalska B, Kowalski D. Application of multidimensional scaling method to identify the factors influencing on reliability of deep wells. In: Burduk A, Chlebus E, Nowakowski T, Tubis A. (eds) Intelligent Systems in Production Engineering and Maintenance. ISPEM 2018. Advances in Intelligent Systems and Computing 2019; 835: 56-65, https://doi.org/10.1007/978-3-319-97490-3_6.
23. Lehner JT, Brown CA. Wear of abrasive particles in slurry during lapping. In: Proceedings of the ASME2014 International Manufacturing Science and Engineering Conference, MSEC 2014 Collocated with the JSME2014, International Conference on Materials and Processing and the 42nd North American Manufacturing Research Conference 2014; 45813: V002T02A01, https://doi.org/10.1115/MSEC2014-4121.
24. Liao BP, Sun B, Li Y, et al. Sealing reliability modeling of aviation seal based on interval uncertainty method and multidimensional response surface. Chinese Journal of Aeronautics 2019; 32(9): 2188-2198, https://doi.org/10.1016/j.cja.2019.01.019.
25. Liao BP, Sun B, Yan MC, et al. Time-variant reliability analysis for rubber O-ring seal considering both material degradation and random load. Materials 2017; 10(10): 1211, https://doi.org/10.3390/ma10101211.
26. Liu D, Wang S, Tomovic MM. Degradation modeling method for rotary lip seal based on failure mechanism analysis and stochastic process. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (3): 381-390, https://doi.org/10.17531/ein.2020.3.1.
27. Loska A, Moczulski W, Wyczółkowski R, Dąbrowski A. Integrated system of control and management of exploitation of water supply system. Diagnostyka 2016; 17(1): 65-74.
28. Loska, A, Paszkowski W. SmartMaintenance - The Concept of Supporting the Exploitation Decision-Making Process in the Selected Technical Network System. In: Burduk A., Mazurkiewicz D. (eds) Intelligent Systems in Production Engineering and Maintenance - ISPEM 2017. Advances in Intelligent Systems and Computing, 2018; 637: 64-73, https://doi.org/10.1007/978-3-319-64465-3_7.
29. Marquez AC. The maintenance management Framework. Models and methods for complex systems maintenance. London: Springer Verlag, 2007.
30. Nicolas L, Pierre B, Cécile M, Nicolas G. AK-MCSi: A Kriging-based method to deal with small failure probabilities and time-consuming models. Structure Safety 2018; 73:1-11, https://doi.org/10.1016/j.strusafe.2018.01.002.
31. Park JY, Han S W, Jeong HS, et al. Advanced sealing system to prevent leakage in hydroforming. Journal of Materials Processing Technology 2017; 247: 103-110, https://doi.org/10.1016/j.jmatprotec.2017.04.006.
32. Shahraki AF, Yadav OP, Liao H. A review on degradation modelling and its engineering applications. International Journal of Performability Engineering 2017; 13(3): 299-314, https://doi.org/10.23940/ijpe.17.03.p6.299314.
33. Sun B, Jiang X, Yung KC, et al. A review of prognostic techniques for high-power white LEDs. IEEE Transactions on Power Electronics 2016; 32(8): 6338-6362, https://doi.org/10.1109/TPEL.2016.2618422.
34. Sun B, Li MM, Liao BP, et al. Time-variant reliability modeling based on hybrid non-probability method. Archive of Applied Mechanics 2020; 90(2): 209-219, https://doi.org/10.1007/s00419-019-01605-1.
35. Sun B, Li Y, Wang Z, et al. An improved inverse Gaussian process with random effects and measurement errors for RUL prediction of hydraulic piston pump. Measurement 2020:108604, https://doi.org/10.1016/j.measurement.2020.108604.
36. Sun B, Yan MC, Feng Q, et al. Gamma degradation process and accelerated model combined reliability analysis method for rubber O-rings. IEEE Access 2018; 6: 10581-10590, https://doi.org/10.1109/ACCESS.2018.2799853.
37. Turner CJ, Emmanouilidis C, Tomiyama T, Tiwari A, Roy R. Intelligent decision support for maintenance: an overview and future trends. International Journal of Computer Integrated Manufacturing 2019; 32: 936-959, https://doi.org/10.1080/0951192X.2019.1667033.
38. Wang L, Wang X, Chen X, et al. Time-variant reliability model and its measure index of structures based on a non-probabilistic interval process. Acta Mechanica 2015; 226(10): 3221-3241, https://doi.org/10.1007/s00707-015-1379-2.

39. Wang P, Yang L, Zhao N, Li L, Wang DA. New SORM Method for Structural Reliability with Hybrid Uncertain Variables. Applied Science,2021, 11: 346, https://doi.org/10.3390/app11010346.
40. Wang P, Zhang J, Zhai H, Qiu, J. A new structural reliability index based on uncertainty theory. Chinese Journal of Aeronautics 2017; 30: 1451-1458, https://doi.org/10.1016/j.cja.2017.04.008.
41. Wronkowicz-Katunin A, Mihaylov G, Dragan K, Timofiejczuk A. Uncertainty Estimation for Ultrasonic Inspection of Composite Aerial Structures. Journal of Nondestructive Evaluation 2019; 38(3): 1-20, https://doi.org /10.1007/s10921-019-0622-5.
42. Yang L, Nie S, Zhang A. Non-probabilistic wear reliability analysis of swash-plate/slipper of water hydraulic piston motor based on convex model. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 2013; 227(3): 609-619, https://doi.org/10.1177/0954406212463501.
43. Yang YJ, Peng WW, Meng D, et al. Reliability analysis of direct drive electrohydraulic servo valves based on a wear degradation process and individual differences. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2014; 228(6): 621-630, https://doi.org/10.1177/1748006X14541256.
44. Yang YJ, Peng WW, Zhu SP, et al. A Bayesian approach for sealing failure analysis considering the non-competing relationship of multiple degradation processes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (1): 10-15, http://doi.org/10.17531/ein.2016.1.2.
45. You L, Zhang J, Li Q, et al. Structural reliability analysis based on fuzzy random uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (4): 599-609, http://doi.org/10.17531/ein.2019.4.9.
46. Zaman K, Rangavajhala S, Mcdonald MP, Mahadevan S. A probabilistic approach for representation of interval uncertainty. Reliability Engineering & System Safety 2011; 96: 117-130, https://doi.org/10.1016/j.ress.2010.07.012.
47. Zhang J, Qiu J, Wang P. Hybrid Reliability Analysis for Spacecraft Docking Lock with Random and Interval Uncertainty. International Journal of Aerospace Engineering 2017: 3920267, https://doi.org/10.1155/2017/3920267.

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Bolesław Karwat, Piotr Rubacha, Emil Stańczyk. Optimization of a screw conveyor's exploitation parameters. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 285–293, http://doi.org/10.17531/ein.2021.2.8.

Optimization of a screw conveyor's exploitation parameters

Indexed by: Web of Science Group

Bolesław Karwat[a], Piotr Rubacha[a], Emil Stańczyk[a]

[a] AGH University of Science And Technology, al. Mickiewicza 30, 30-059 Kraków

Highlights

- Verification of theoretical design methods with experiments and simulations results.
- Computer DEM simulations as an advanced tool for designing screw conveyors.
- Optimization of construction and exploitation parameters of a screw conveyor.

Abstract

The paper describes the problem of designing screw conveyors in terms of determining their exploitation characteristics. Based on the actual values of mass efficiency and power demand obtained in a laboratory experiment, the theoretical design methods and the numerical discrete element method model results were verified. The obtained results have shown that the currently used theoretical methods underestimate the mass efficiency and power demand compared to experiments when typical values of filling rate coefficient and progress resistance coefficient are used. It was also shown that the results of DEM simulations are in good agreement with the experiments in terms of mass efficiency and power demand. Based on the exploitation characteristics determined in DEM simulations for different constructions of the screw and different rotational speeds, multi-objective optimization of the exploitation parameters of the screw was performed in order to minimize the power demand of a screw conveyor and simultaneously maximize its mass efficiency. The optimization results showed that it is possible to find such construction and the rotational speed that will maximize the mass efficiency of the conveyor and keep the power demand low, reducing the exploitation costs of the device.

Keywords

screw conveyor, discrete element method, bulk material, exploitation characteristics.

1. Introduction

Screw conveyors are intended for the short-distance transportation of bulk materials in food, agricultural, energy, or lime industry plants. They are also used when other processes, such as mixing, heating, cooling, drying, moistening, or precise proportioning, are required. A screw conveyor can be used for transportation in three planes: vertical, horizontal, or at a certain inclination angle. Simple construction and working principle, together with an easy adjustment of their efficiency through the change of the shaft's rotational speed, are the key advantages of the screw conveyors. Currently, screw conveyors can be equipped with screw flights of diverse shapes. Depending on the application, the screws may have varying pitches, conical internal and external diameters through the whole length, and the flights may be irregular in shape. In so-called reversible conveyors, used for bi-directional transportation of material, we can even find flights of opposite twist directions mounted on one shaft. Such untypical constructions of screws are possible with advanced flight production technologies, such as forming in hydraulic presses without preheating the prefabricated parts.

Producing any shapes and sizes of the screw flights is not a technological problem. What is challenging for the engineers is the designing process of the screw conveyors. Their construction and exploitation parameters must meet the defined requirements concerning the mass efficiency, filling rate of the trough, or providing the necessary power to the drive. The behavior of bulk materials during transportation by a screw conveyor is very complicated. It depends on many factors, such as the type and shape of the screw flights, the rotational speed of the shaft, the way of proportioning of the material, or the physical properties of the material. Theoretical methods for designing the screw conveyors do not consider all the factors mentioned above or oversimplify them. In the case of typical bulk materials of uniform granulation and standard constructions of the screws (constant pitch, constant internal and external diameters) and fed by one source, e.g., a hopper, theoretical methods allow reasonable estimation of the exploitation parameters of the screw conveyors. In the case of materials of specific properties (cohesive or strongly aerated materials) or for unusual shapes of screws or multiple feeding points, these methods do not provide reliable results of efficiency and power [16]. For this reason, the external diameters of the screws and power demands are very often chosen to be safely larger. Such an approach is unfavorable because of the excessive use of materials, ineffective use of the drive unit, and high exploitation costs. On the other hand, the difficulties in making decisions on the construction of screw conveyors are caused

E-mail addresses: B. Karwat - karwat@agh.edu.pl, P. Rubacha - prubacha@agh.edu.pl, E. Stańczyk - stanczykemil@gmail.com

by insufficient exploitation data of the existing systems. Implementing the diagnostic, measuring, and expert systems would allow collecting and processing of the data on the exploitation parameters of the screw conveyors, which are working under different conditions in the industry [10, 12-14].

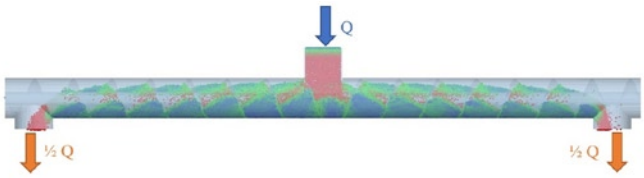


Fig. 1. A DEM simulation of a reversible conveyor

Nowadays, the design can be aided by advanced tools of virtual prototyping [4, 13, 17, 28]. Computer simulations are also used for studying the behavior of bulk materials during transportation by a screw conveyor. High computational powers of the workstations and highly developed numerical methods allow an accurate representation of the physical phenomena. As a result, the physical properties of a given bulk material and its interactions with a conveyor's components can be considered. In the simulations of bulk materials' behavior, the discrete element method (DEM) is used. Numerous simulation studies have confirmed the reliability of the results obtained in the simulations with comparison to the actual measurement results for different machines and devices used for the transportation of bulk materials [9, 24-25].

Screw conveyors have been a subject of research for years, studied theoretically, experimentally, and in computer simulations. The studies published so far have identified several phenomena observed during the transportation of a material by a screw conveyor. This, in turn, was used to determine how the physical properties of the transported materials, the geometry of the screw, the rotational speed of the shaft, and the conditions of material feeding influence the exploitation parameters of the device.

The authors of [15, 20-21, 25] have shown a good agreement between the results of DEM simulations and the results of experiments in determining a screw conveyor's efficiency. Besides, the research in [20] shows that theoretical models overestimate the values of mass efficiency compared to the actual results. The authors of [9] have shown a good agreement between the results of DEM simulations and the experiments in determining power demand. In [19], the authors used DEM to investigate the influence of filling rate, the rotational speed of the screw, the external diameter of the shaft, and the external friction coefficient of the material on the efficiency of a screw conveyor. The paper [11] shows the methodology for the calibration of a DEM material model during the simulation of the transportation of bulk materials by the screw conveyors of varying geometries (a screw with a constant pitch and external diameter on a regular shaft, a screw with changing external diameter, a screw with a varying pitch, a screw with a shaft of a changing external diameter). A good agreement was obtained between the efficiencies obtained in the simulations and experiments for all the studied variations of the screws, except for the screw with varying pitch – in this case, the simulated mass efficiency was underestimated by 24%. The authors explain this discrepancy with the size of DEM particles; they were too large compared to the screw pitch. In [27], the influence of the conditions of feeding the bulk material to the conveyor on the mass efficiency and power demand of the device was studied in laboratory conditions. It was shown that the efficiency of a screw conveyor increases with the rotational speed of a shaft, up to a value above which the centrifugal force causes the movement of the material towards the outside, which as a result limits the metering of the material from the hopper to the inside of the trough. It was also observed that the efficiency of the conveyor and its power demand are determined by the size of the trough inlet opening. Higher efficiency values were obtained for larger inlet openings. What is more, the paper [18] showed that the largest possible inlet opening together with a low rotational speed of the shaft provided the best efficiency because of the minimized power demand. The authors of [26] showed a correlation between the physical properties of the bulk material and the efficiency of the screw conveyor, based on the performed laboratory tests (shear cell test, compressibility test, permeability test, dynamic flow test). The papers [2,22,29] showed discrepancies between the experimental and theoretical efficiencies of a tubular screw conveyor. According to the theoretical models, the efficiency increases linearly with the increase of the rotational speed of the screw. In reality, the efficiency increases to a specific limit value of the rotational speed above which the efficiency curve flattens. It means that the further increase of rotational speed causes a smaller increase in efficiency. Paper [3] shows the results of the experiments on the power demand of a screw conveyor. It was observed that in the case of not aerated materials (sand and gravel), the power demand increases proportionally to the rotational speed of the screw. The increase of the rotational speed caused the decrease of the power demand for the transportation of an aerated material. It was caused by the decreased bulk density of the material after aeration. The authors of [23], based on the experiments' results, concluded that the smaller the ratio of the screw pitch and its external diameter, the greater the mass efficiency. The authors suggested that there is such a construction of a conveyor that ensures effective relation between efficiency and power demand. A good agreement between theoretically and experimentally obtained values of mass efficiency and power demand of a conveyor was observed. The calculations considered the working conditions and physical properties of the transported material (sand) very accurately, which, according to the authors, allowed such a good agreement.

Many papers correctly and purposefully indicated the influence of several factors connected with the transported material's properties, ways of manufacturing the material, or the construction and exploitation parameters of the screw on its work parameters. However, there are no clear guidelines for the constructors on how to include these factors in the designs using theoretical methods. The authors of some of the papers have confirmed the weaknesses of the theoretical methods for calculating the screw conveyors. The up-to-date simulation studies have shown a very good agreement between DEM simulations and experimental results regarding the mass efficiency of a screw conveyor. However, little attention is paid to the estimation of the power demand of a screw conveyor using the DEM method. Correct determination of power demand and efficiency of a screw conveyor allows the determination of exploitation characteristics, facilitating the choice of construction parameters and rotational speed to minimize the power demand and ensure effective transportation. There is also a lack of papers on the possibilities of optimizing the construction of the screw conveyors. This is why this paper shows the DEM simulation results of the transportation of cement performed to estimate the mass efficiency and power demand based on the experimental results. What is more, the results of multi-objective optimization of the exploitation parameters of a conveyor aimed at minimizing the power demand and maximizing the efficiency of a conveyor, performed with the use of the DEM method, are presented.

2. Methods of determining the exploitation parameters of a screw conveyor

The currently used theoretical methods allow the determination of mass efficiency of a conveyor and the drive power based on the adopted dimensions of the screw (pitch, external diameter), bulk density of the material, the rotational speed of the shaft, filling rate, and the progress resistance coefficient [1,16]. The basic method of determining a conveyor's mass efficiency and its power demand is described by Eq. (1) and (2).

$$Q = 60 \cdot \frac{\pi \cdot D_z^2}{4} \cdot n \cdot s \cdot \Psi \cdot \rho \cdot k_n, t/h \quad (1)$$

where:

- D_z – external diameter of the screw, m,
- n – rotational speed of the screw, rpm,
- s – screw pitch, m,
- ρ – bulk density of the material, t/m^3,
- ψ – filling rate (within the range of 0.15-0.45, depending on the type of the material),
- k_n – coefficient dependent on the angle of inclination of the conveyor (for a horizontal conveyor, the coefficient takes the value of 1),

$$P_H = \frac{c_o \cdot Q \cdot L}{367}, kW \quad (2)$$

where:

- c_o – progress resistance coefficient of the transported material, dependent on the type of the material,
- Q – efficiency of the conveyor, t/h,
- L – length of the conveyor, m.

Eq. (3), proposed by CEMA (Conveyer Equipment Manufacturers Association), is very similar to Eq. (1), defined as per the basic method, i.e., mass efficiency is the function of the dimensions of the screw, rotational speed, and the filling rate of the trough. The CEMA method uses the imperial units, and the calculation models are shown in such a system. All the quantity units are taken according to CEMA [1]:

$$C = \frac{0.7854 \cdot \left(D_s^2 - D_p^2\right) \cdot P \cdot K \cdot 60 \cdot n \cdot W}{1728}, lbs/h \quad (3)$$

where:

- C – mass efficiency, ft^3/h,
- n – rotational speed of the shaft, rpm,
- D_s – external diameter of the screw, in,
- D_p – diameter of the shaft of the screw, in,
- P – screw pitch, in,
- K – filling rate percent,
- W – bulk density of the material, lbs/ ft^3.

Eq. (4) expresses the power required for the transportation of material by a screw conveyor, without including the efficiency of the drive and friction in the bearings. Unlike the basic method, this equation includes the type and the shape of the screw flights.

$$hp_m = \frac{C \cdot L \cdot W \cdot N \cdot F_f \cdot F_m \cdot F_p}{1 \cdot 10^6}, hp \quad (4)$$

where:

- C – mass efficiency, ft^3/h,
- L – length of the conveyor, ft,
- W – bulk density of the material, lbs/ft^3,
- F_f – flight factor,
- F_m – material factor,
- F_p – paddle factor.

The equations for mass efficiency and power demand shown above are linear relations of the screw dimensions, filling rate of the trough, bulk density of the material, and empirical progress resistance coefficient. However, because of the limitations o these methods, the calculated exploitation parameters are not reliable. With an inaccurate calculation method, it is impossible to optimize the construction and exploitation parameters of a screw conveyor because of the significant risk of underestimating the results.

Unlike the theoretical methods, the numerical discrete element method allows reliable simulations of the bulk materials' behavior since it accounts for the interactions between the grains of the modeled material and the interactions between the material and the parts of the device. The modeled material is represented by a number of rigid spheres, whose behavior is described by various contact models [7, 9]. These models simulate the behavior of different materials, such as dry materials, strongly cohesive materials, or highly compressive materials. A DEM simulation accounts for the physical properties of the bulk material and, as a result, reflects its behavior very reliably. The material DEM model's basic inputs are the shape and size of a particle representing the actual grain of the material, the density of a particle, external and internal friction coefficients, and the restitution coefficient. It is worth noting that the material model's parameters in a DEM simulation are defined microscopically, i.e., for individual particles. Therefore, to reflect the macroscopic physical properties, the properties of individual particles must be chosen in a way that enables reflecting the behavior of the whole material. This process is called the DEM material model calibration and is further described by [5, 6, 7]. Having calibrated the model, we can simulate the transportation of the bulk material by a screw conveyor, determine its exploitation parameters (mass efficiency, power demand), and assess the filling rate of the trough, the behavior of the material in the feeding zone, and the abrasion wear of the flights.

3. Experimental and simulation studies on the exploitation parameters of a screw conveyor

The experiment aimed to measure the actual exploitation parameters of a screw conveyor, such as mass efficiency and power demand, depending on the rotational speed of the screw. The experiment was performed on a laboratory line designed for determining the exploitation parameters of a screw conveyor. The visualization of the laboratory line is shown in Fig. 2.

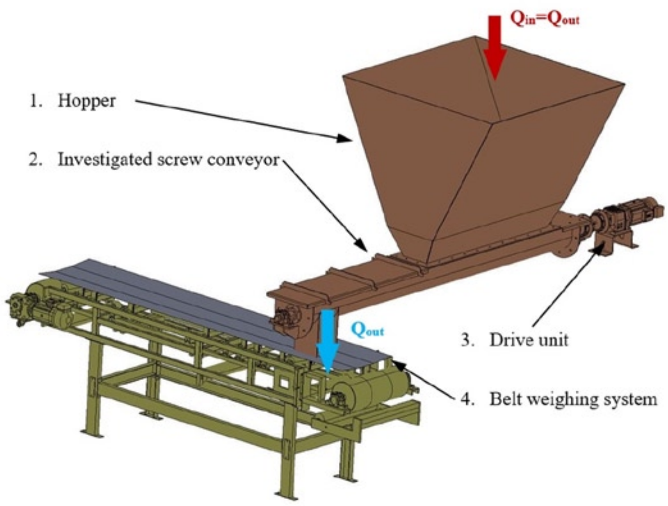


Fig. 2. Visualization of the laboratory line.

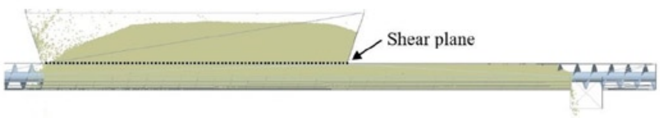
The experimental laboratory line was equipped with a hopper of 4 m^3 volume. The screw conveyor under study, with an external diameter of 160 mm and the screw pitch of 75 mm, was mounted on a shaft of a 70 mm diameter. A helical gear unit with a frequency converter supplied the power. The material from the screw conveyor was transported to a weighing unit type WMTP B-650, which allowed the measurement of the efficiency within the range up to 10 Mg/h, with an accuracy of 0.5%. The material from the weighting unit was then transported to a belt conveyor that fed the elevator transporting the material back to the hopper and closing the loop. The power demand was measured as the power consumption of the drive unit at a given rotational speed. Dimensions and exploitation parameters of the screw conveyor under study are shown in Table 1.

Table 1. Dimensions and exploitation parameters of the screw conveyor under study.

Parameter	Value
D_z – external diameter of the screw	160 mm
d_w – diameter of the screw shaft	70 mm
P – screw pitch	75 mm
L – length of the conveyor	4000 mm
n – the range of the rotational speed	20 – 70 rpm
α – the angle of inclination of the conveyor	0°

In order to minimize the increase of the power demand caused by the shear in the layer of the material at the interface of the hopper and the trough, the amount of the material was chosen to fill the trough without accumulating in the hopper as shown in Fig. 3.

a)



b)

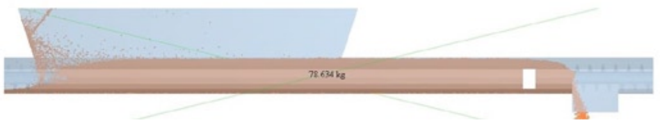


Fig. 3. a) Shear layer at the interface of the hopper and the trough b) Filling rate of the trough during the experiment

The efficiencies of the belt conveyors and the bucket conveyor used for the closed-loop of the transported material were chosen to fill the trough, without the aforementioned accumulation of the material in the hopper. A Portland cement was used in the experiment, as it is a representative bulk material used in the cement, energy, and chemical industry plants. The granulation of the Portland cement is very fine, within 1μm to 50 μm. Its physical properties are very similar to the materials such as gypsum, limestone powder, and raw material powder. In the experiment and in the simulations, the assumption was made that the transported material is dry and non-cohesive. The physical properties of the Portland cement are shown in Table 2.

Table 2. Physical properties of the Portland cement

Physical property	Value
Bulk density, kg/m^3	1050
Moisture, %	0.2
The angle of repose, °	35.9
External friction coefficient	0.52

The DEM simulation model of the screw conveyor reflected the actual object used for the experiment. In order to ensure similar working conditions, the model included the hopper, the trough, and the central working unit – the screw, as shown in Fig. 4.

Only the screw conveyor and the hopper were included in the simulations, without the remaining equipment of the closed transportation loop. The constant amount of the material was ensured by the use of the periodic boundary condition. It means that the material from the end of the trough was fed back to the hopper, creating the closed loop. The parameters of the DEM model are shown in Table 3.

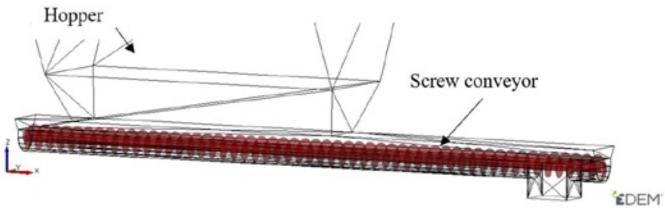


Fig. 4. The geometrical model under study

Table 3. Parameters of the material model of the cement used in the DEM simulations

Parameter	Value
Size of a DEM particle	3 mm
The shape of a DEM particle	Sphere
Shear modulus (G) for the interaction between the particles	1.0e+7 MPa
Shear modulus (G) for the interactions between the particles and the walls	1.0e+7 MPa
Poisson ratio	0.25
The density of a DEM particle	1675 kg/m^3
Coefficient of restitution	0.5
Internal friction coefficient	0.25
Rolling resistance coefficient between the particles	0.1
External friction coefficient	0.52
Rolling resistance coefficient between the particles and the walls	0.01
Timestep	2.66e-5 s
Number of particles used in the simulation	422, 000

4. Results

Fig. 5 shows the mass efficiency as a function of the rotational speed of the screw shaft, obtained in experiments, simulations, and using the theoretical models. In the calculations, typical values of the filling rate of the trough were adopted. In the basic theoretical model, the filling rate was set to $\psi = 0.25$ and the inclination coefficient $k_n = 1.0$. In the CEMA method, the filling rate coefficient was set to

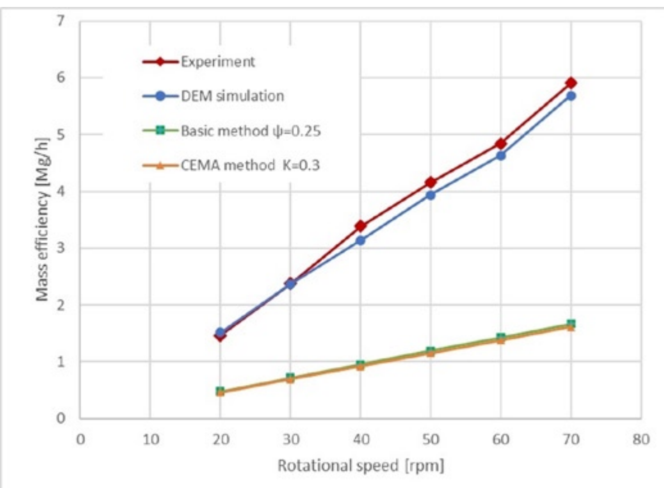


Fig. 5. Mass efficiency as a function of the rotational speed of a screw shaft

K = 0.3. The results obtained in the CEMA method were converted from the imperial units to the SI units.

In Fig. 5 we can see a good agreement between the DEM simulations and the experiment results. The efficiencies obtained in the basic theoretical method and CEMA are also very similar. However, with the adopted values of the filling rate, the theoretical values are much

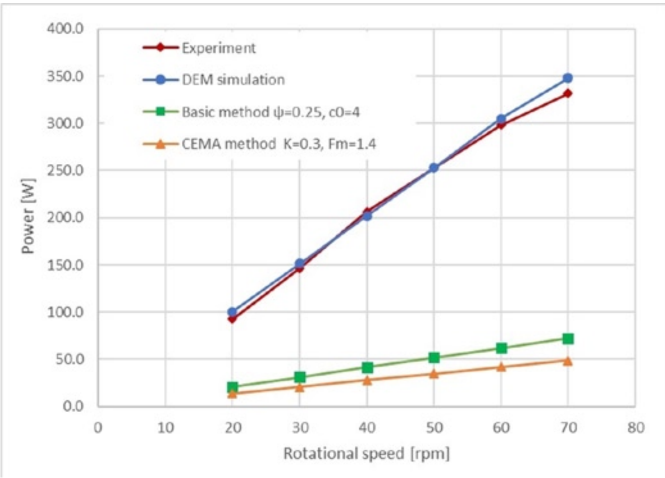


Fig. 6. Power demand as a function of the rotational speed of the screw shaf

lower than those obtained in the DEM simulations and the experiment.

Fig. 6 shows the power demand as a function of the rotational speed of the screw shaft in a similar way. The actual power required for the transportation of the cement with a defined rotational speed was determined as a difference of the total power measured during the transportation of the material and the power of an empty conveyor. For the theoretical calculations, in the Portland cement case, the resistance coefficients were taken to be $c_0 = 4.0$ for the basic method and $F_m = 1.4$ for the CEMA method.

Fig. 6 shows that the power demand of a crew conveyor during the cement's transportation determined in DEM simulations is in good agreement with the actual results of the experiment. The theoretical calculations underestimated the results significantly compared to experiments and DEM simulations. What is more, there is a discrepancy between the results of the basic and CEMA methods. For the adopted values of the material's progress resistance coefficient inside the trough, the CEMA method resulted in lower values.

The use of DEM simulations facilitated the assessment of the filling rate of the trough at different values of the rotational speed. Fig. 7 shows the filling rate of the trough at the rotational speed equal to 20, 50, and 70 rpm.

The presented distributions of the filling rates of the trough show that the trough is filled in 100% through the whole length of the screw conveyor. Based on the above results, the filling rate coefficients were corrected to $\psi = 0.9$ for the basic method and K = 1.0 for the CEMA method, which significantly improved the consistency between the theory and the experiment. Fig. 8 shows the comparison of the results of mass efficiency for the corrected filling rate coefficients.

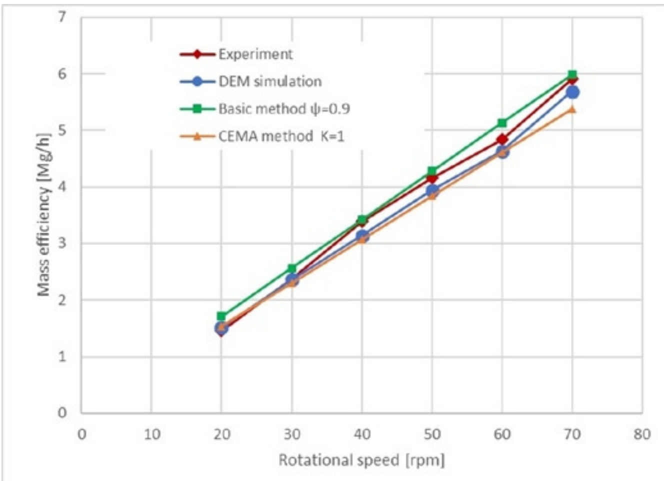


Fig. 8. Mass efficiency results for the corrected values of filling rate coefficients

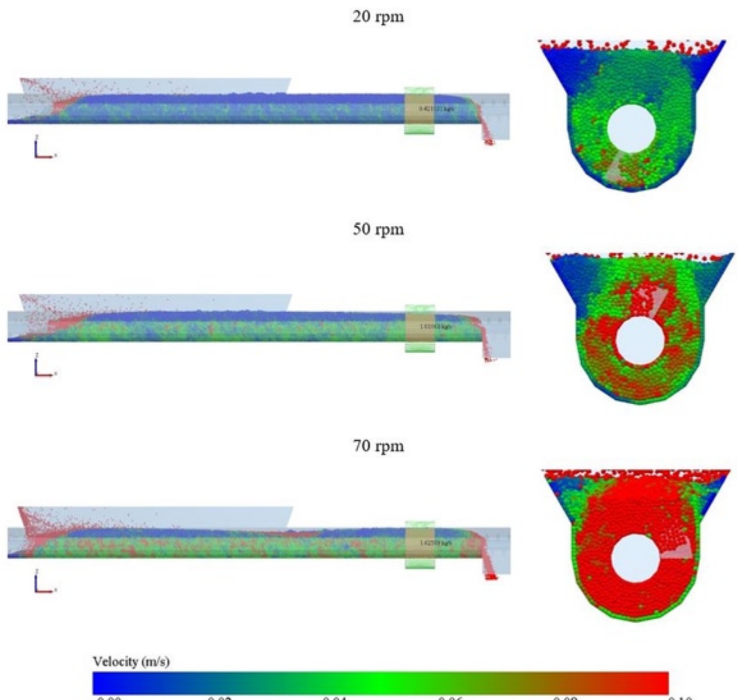


Fig. 7. Filling rate of the trough together with the distribution of the velocity of the material in the cross-section through the trough

With the corrected values of filling rate coefficients both theoretical methods resulted in values very similar to the results of the DEM simulations and the experiment.

A similar procedure was repeated to determine the power required for the transportation of the material by a screw conveyor. The mass efficiencies calculated for the corrected values of the filling rate coefficients were used, and the progress resistance coefficients were chosen in such a way that allowed a good agreement with the actual results of experiments. For the basic method, the coefficient c_0 was determined to be $c_0 = 5.5$, and for the CEMA method, the coefficient $F_m = 3.1$. Fig. 9 shows the comparison of the power demand with the corrected resistance coefficients.

Fig. 9 implies that using the corrected progress resistance coefficients results in a good agreement between the results obtained during the laboratory tests and determined in the DEM simulations.

5. Multi-objective optimization

Decision-making is an integral part of the design process. The constructors decide which solutions are the most effective for specific criteria, such as efficiency, power, durability, etc. Usually, these criteria (objectives) are conflicting, e.g., designing a durable device with a minimized use of materials, achieving the best possible efficiency with minimum power demand, etc. In most cases, such solutions cannot be found. This is because the minimization of one objective causes the maximization of the other. We can say that the objectives are conflicting, but one dominates the other [7]. Hence, the optimization results can be divided into dominated and non-dominated solutions.

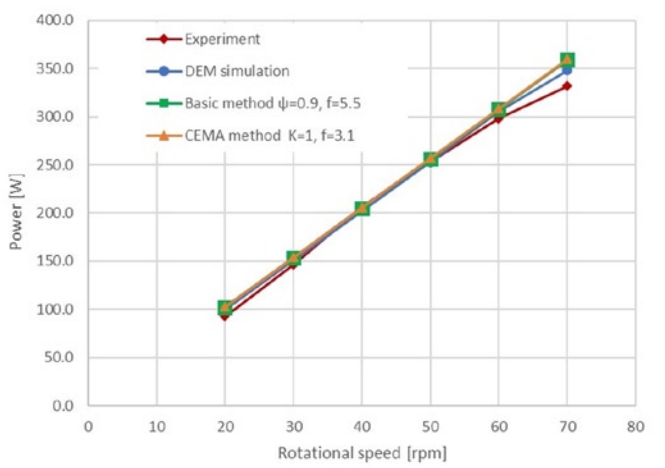


Fig. 9. The power demand for the corrected values of the resistance coefficients as a function of the rotational speed of the screw

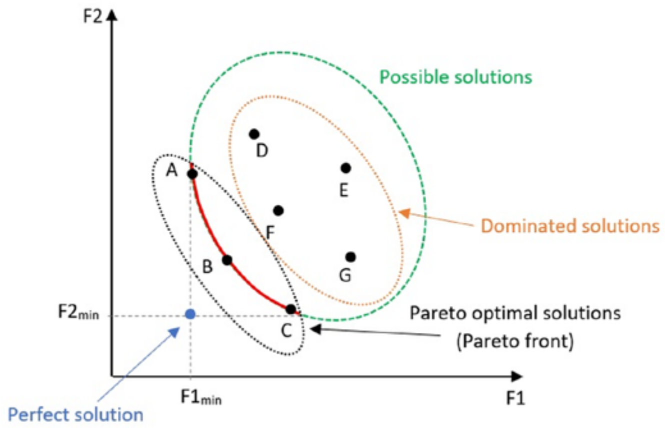


Fig. 10. Sets of dominated and non-dominated solutions

The dominated solutions are the ones, for which both criteria can be improved. In Fig. 10 these are the solutions D, E, F, and G. The choice of such solutions is unjustified since there are better solutions in terms of $F1_{min}$ and $F2_{min}$. Such solutions are called non-dominated or Pareto optimal [7]. This means that a solution that improves one of the criteria without degrading the other does not exist. Figure 10 shows a set of solutions A, B, and C called a Pareto front; the elements of such a set are Pareto optimal. A perfect solution is such a vector of variables, for which all the objective functions take the minimum value. Finding such a point is practically very rare since each objective function takes its minimum for a different set of variables [7].

In the case of the above example, we can specify three variants. Solution A reaches the minimum of the objective function F1 and the maximum of the objective F2, Pareto optimally. Solution B takes middle values of F1 and F2, and solution C reaches the minimum of F2 and the maximum of F1, Pareto optimally. From this example, we can see that in multi-objective optimization, the decision-maker can choose a solution from the Pareto front depending on their preferences.

Mathematically, a general decision-making model takes the following form:

$$F(x) = f(x_1, \ldots, x_n) \rightarrow min, max$$
$$g_i(x_1, \ldots, x_n) \geq 0, i = 1, \ldots, m \tag{5}$$

where:

F(x) – objective function (criterion), which should be minimized or maximized,

$x_1, \ldots x_n$ – variables,

$g(x_1, \ldots x_n)$ – constraints.

The value of the objective function (criterion) can be defined as a measure of the decision to be made. Depending on the requirements, this function can be either minimized or maximized. The variables describe alternative decisions. In many cases, the decision must be made considering numerous criteria F1(x), F2(x), …, which comes to the problem of multi-objective optimization.

6. Results of the multi-objective optimization of the exploitation parameters of a screw conveyor

Based on the calibrated DEM material model, several simulations were performed for different constructions of a screw conveyor and different rotational speeds in order to determine the efficiency characteristics and power demand. Table 4 shows the construction parameters of the screw conveyors under study.

Table 4. Construction details of the simulated screw conveyors.

Construction No.	D – external diameter, mm	d – internal diameter, mm	s – screw pitch, mm
1	150	60	100
2	150	60	150
3	150	60	200
4	200	60	100
5	200	60	150
6	200	60	200

The simulations were performed for the rotational speeds within the range of 20 to 80 rpm. In order to ensure identical working conditions in each case, the amount of material in the hopper and the trough was constant and equal to 150 kg, as shown in Fig. 11.

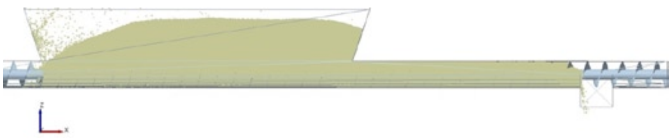


Fig. 11. Simulated transportation of the material by a screw conveyor

Fig. 12 shows the simulated mass efficiency values for each construction of the conveyor as a function of the rotational speed of the screw.

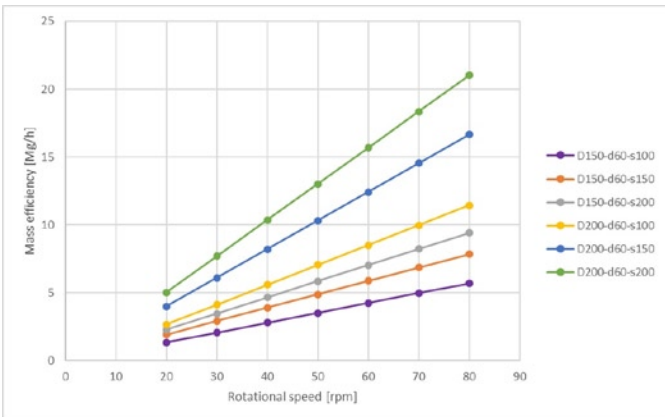


Fig. 12. Mass efficiency of the simulated screw conveyors as a function of the rotational speed of the screw shaft

As it is indicated by the above graphs, the efficiency increases linearly within the studied range of the rotational speed of the shaft. For individual rotational speeds, the efficiency increases with the increase of the screw pitch and the flights' external diameter.

Fig. 13 shows the power demand required for cement transportation by a screw conveyor at defined rotational speeds of the screw shaft.

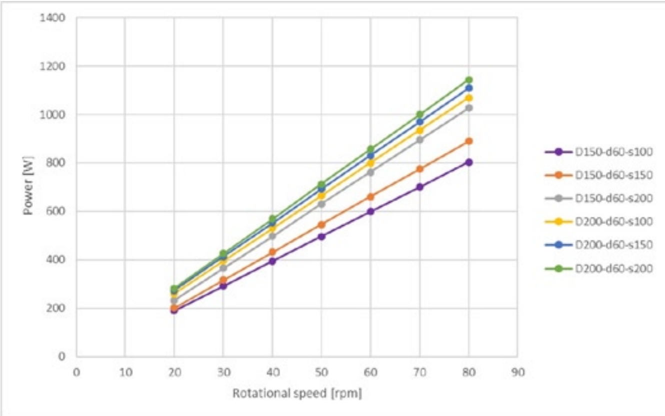


Fig. 13. Power demand of the simulated screw conveyors as a function of the rotational speed of the screw shaft

Similarly, like in the case of mass efficiency, the power demand of each conveyor increases together with the increase of the rotational speed. We can see that also in this case at a constant rotational speed, the power demand increases with the increase of the screw pitch and the flights' external diameter.

Both characteristics show how the construction parameters of the screw conveyors and the rotational speed of the shaft influence the mass efficiency and power demand. The interpretation of the results does not determine which option is the best possible to maximize the efficiency with simultaneous minimization of the power demand. Therefore, in order to be able to choose the best construction, we can perform multi-objective optimization.

The exploitation parameters of a screw conveyor, i.e., mass efficiency and power demand are determined by the construction parameters, such as the external diameter of a screw and screw pitch. According to the principles of multi-objective optimization, the following objective functions were determined:

F1 – power required for the transportation of the material (6)

$$F1(x) = f(x_1, x_2, x_3) \rightarrow min \tag{6}$$

F2 – mass efficiency of a conveyor (7)

$$F2(x) = f(x_1, x_2, x_3) \rightarrow max \tag{7}$$

With such objective functions, the following variables were taken into account:

x_1 – external diameter of the screw,

x_2 – screw pitch,

x_3 – the rotational speed of the screw.

The constraints g(x) were adopted as follows: rotational speed within 40-60 rpm, screw pitch within 100-200 mm, and the external diameter of the screw flights within 150-200 mm. This relation is described by Eq. (8):

$$g(x) = f(x_1, x_2, x_3) \tag{8}$$

For such an optimization task, the optimization of a screw conveyor's construction parameters can be performed. Table 5 shows the defined objective functions and constraints.

Table 5. Defined objective functions and constraints.

F1(x)	**F2(x)**	**g(x)**
Power demand → minimum	Efficiency → maximum	$g(x_1) \in$ <100; 200 mm> $g(x_2) \in$ <150; 200 mm> $g(x_3) \in$ <40; 60 rpm>

Figure 14 shows the possible solutions in the space of objective functions, together with a perfect solution.

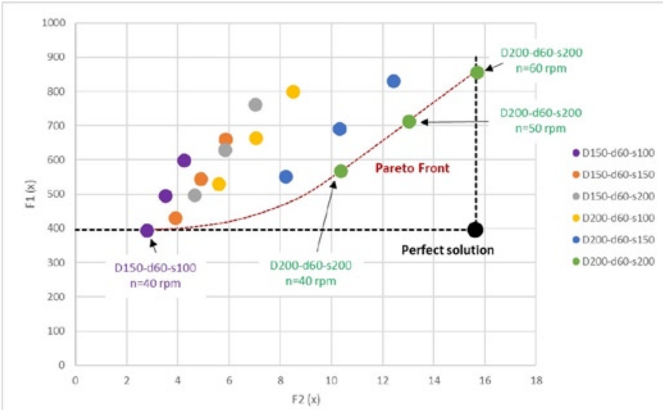


Fig. 14. Possible solutions of the optimization and a Pareto front

The Pareto front was obtained for three values of the rotational speed (40, 50, and 60 rpm), the external diameter of the conveyor D_{out}=200 mm and the screw pitch s=200 mm, and for the rotational speed equal to 40 rpm for a screw conveyor of the external diameter of D_{out}=150 mm and the screw pitch s=100 mm. Obtaining a perfect solution in terms of (6) and (7) is impossible due to the adopted constraints. As can be easily seen, these functions are conflicting, and the increase in efficiency caused the increase of the power demand.

When one of the objective functions is minimized, and the other one is maximized, the optimal solutions of (6) and (7) can be found by solving Eq. (9):

$$\max\left(\frac{F1(x)}{F2(x)} : x \in D\right) \tag{9}$$

where:

D – the set of possible solutions,

$F2(x) \neq 0 : x \in D$

For each construction of a screw conveyor, the ratios of F1(x) – power demand and F2(x) – mass efficiency were determined. The values of the obtained ratios are shown in Table 6.

Table 6. Ratios of the objective functions

Construction details	F1(x)/F2(x)
D150-d60-s100	0.0071
D150-d60-s150	0.0090
D150-d60-s200	0.0093
D200-d60-s100	0.0106
D200-d60-s150	0.0149
D200-d60-s200	0.0183

The maximum ratio of the objective functions F1(x) and F2(x) was obtained for the screw conveyor of the external diameter D_{out}=200 mm and the screw pitch s=200 mm. It means that this construction is the optimal solution of (6) and (7) with the adopted constraints.

7. Conclusions

Designing screw conveyors requires much attention and caution from the constructors. The conveyors' construction parameters (external diameter, screw pitch) and the rotational speed of the screw significantly influence the exploitation parameters, i.e., efficiency and power demand. The goal is to ensure the required mass efficiency at the defined filling rate of the trough and the rotational speed of the screw, and to allow efficient transportation of the material. On the other hand it is important to ensure some capacity reserve in cased a temporary increase of the material fed to the conveyor. As shown in the so-far research and by the experiments' results, theoretical calculation methods are not sufficient since they do not provide reliable results of mass efficiency and power demand required for the transportation of the material. The theoretical methods are based on the assumption that mass efficiency and power demand are the functions of the dimensions of the screw, bulk density, filling rate of the trough, and empirical progress resistance coefficient of the bulk material inside the trough. Such methods simplify the physical properties of the transported materials. They do not include their interaction with the trough and the screw and the bulk density change during the transportation. Keeping in mind all the limitations of these methods and the fact that the results are very often far from reality, the constructors rarely risk optimizing the construction of the screw conveyors.

As shown in the simulations, the results of the mass efficiency ad power demand obtained in the numerical discrete element method (DEM) are in good agreement with the experiments' results. This method reflects the behavior of a bulk material during transportation by a screw conveyor since it includes the physical properties of the materials and their interaction with the components of the screw conveyors. Large possibilities offered by the DEM method for modeling bulk materials make it a universal tool aiding the design of screw conveyors. In order to obtain reliable results of the simulation, the input parameters of the DEM material model must be calibrated carefully, based on the actual physical properties of the bulk material. Computer simulations allow the determination of mass efficiency and power demand of a screw conveyor and the estimation of filling rate of the trough, assessment of the behavior of the bulk material in the feeding zone, and the abrasion wear of the flights. With such a reliable tool, the construction parameters of the screw conveyor can be optimized to decrease its power demand (and, as a consequence, exploitation costs caused by the electric energy consumption) with the maintained high efficiency. As it was shown by the results of the performed multi-objective optimization, the transportation is more effective with a conveyor of a larger external diameter of the flights and larger screw pitch (the efficiency is maximized, and the power demand is minimized). Obtaining a perfect Pareto solution was not possible due to the adopted constraints. Nevertheless, the multi-objective optimization facilitates the decision on the choice of the construction parameters and the rotational speed of the screw.

Further research on the optimization of the construction of screw conveyors should include additional objectives concerning minimization of flights' abrasion wear and minimization of the conveyor's mass. Together with the minimization of the power demand and maximization of the efficiency, these functions may allow such construction that will reduce the costs of producing, exploitation (e.g. replacement of the worn flights), and the power supplied to the device while working.

References

1. ANSI/CEMA standard no. 350. Screw Conveyors for Bulk Materials. 5th edition 2019.
2. Bates L. Entrainment patterns of screw hopper dischargers, transaction of the ASME, Journal of Engineering for Industry 1969; 91: 295-302, https://doi.org/10.1115/1.3591561.
3. Carleton A J, Miles J E P, Valentin F H H. A study of factors affecting the performance of screw conveyors and feeders, Journal of Engineering for Industry 1969; 91: 329-333, https://doi.org/10.1115/1.3591565.
4. Chudzik A, Warda B. Fatigue life prediction of a radial cylindrical roller bearing subjected to a combined load using FEM. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (2): 212-220, https://doi.org/10.17531/ein.2020.2.4.
5. Coetzee C J. Calibration of the Discrete Element Method and the effect of particle shape, Powder Technology 2016; 297: 50-70, https://doi.org/10.1016/j.powtec.2016.04.003.
6. Coetzee C J. Review: Calibration of the discrete element method. Powder Technology 2017; 310: 104-142, https://doi.org/10.1016/j.powtec.2017.01.015.
7. Cunha A G, Covas J A. Optimization in Polymer Processing. Nova Science, New York, United States 2011.
8. Karwat B, Machnik R, Niedźwiedzki J, Nogaj M, Rubacha P, Stańczyk E. Calibration of bulk material model in discrete element method on example of perlite D18-DN, Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 351-357, https://doi.org/10.17531/ein.2019.2.20.
9. Karwat B, Rubacha P, Stańczyk E. Simulational and experimental determination of the exploitation parameters of a screw conveyor. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (4): 741-747, https://doi.org/10.17531/ein.2020.4.18.
10. Kozłowski E, Mazurkiewicz D, Żabiński T, Prucnal S, Sęp J. Machining sensor data management for operation-level predictive model. Expert Systems with Applications 2020; 159: 1-22, https://doi.org/10.1016/j.eswa.2020.113600.
11. Kretz D, Callau-Monje S, Hitschler M, Hien ., Raedle M, Hesser J. Discrete element method (DEM) simulation and validation of a screw feeder system, Powder Technology 2016; 287: 131-138, https://doi.org/10.1016/j.powtec.2015.09.038.
12. Kulinowski P, Kasza P, Zarzycki J. Identification of the operating parameters of the friction drum drive in industrial conditions. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23 (1): 94-102, https://doi.org/10.17531/ein.2021.1.10.
13. Mazurkiewicz D. A Knowledge Base of the Functional Properties of the Conveyor Belt Adhesive Joint for FEM Simulation of Its Stress and Strain State. Journal of Adhesion Science and Technology 2012; 26 (10-11): 1429-1442, https://doi.org/10.1163/156856111X618308.
14. Mazurkiewicz D. Computer-aided maintenance and reliability management systems for conveyor belts. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16 (3): 377-382.
15. McBride W, Cleary P W. An investigation and optimization of the 'OLDS' elevator using Discrete Element Modeling. Powder Technology 2009; 193: 216-234, https://doi.org/10.1016/j.powtec.2009.03.014.
16. Minglani D, Sharma A, Pandey H, Dayal R, Joshi J B, Subramaniam S. A review of granular flow in screw feeders and conveyors, Powder Technology 2020; 366: 369-381, https://doi.org/10.1016/j.powtec.2020.02.066.
17. Mitura A, Gawryluk J, Teter A. Numerical and experimental studies on the rotating rotor with three active composite blades. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (4): 571-579, https://doi.org/10.17531/ein.2017.4.11.
18. O'Callaghan, J R. Some experiments on the intake processing a vertical screw conveyor, Journal of Agricultural Engineering 1962; 7(4): 282-287.

19. Oreficea L, Khinast J G. DEM study of granular transport in partially filled horizontal screw conveyors, Powder Technology 2017; 305: 347-356, https://doi.org/10.1016/j.powtec.2016.09.067.
20. Owen P J, Cleary P W. Prediction of screw conveyor performance using the Discrete Element Method (DEM). Powder Technology 2009; 193: 274-288, https://doi.org/10.1016/j.powtec.2009.03.012.
21. Patinge S, Prasad K. Screw feeder performance prediction using Discrete Element Method (DEM), International Journal of Scientific & Engineering Research 2017; 8(3).
22. Roberts A W. Design and performance criteria for screw conveyors in bulk solids operation, Bulk Solids Handling 2002; 22(6): 436-444.
23. Uematu T, Nakamura S. A study of the screw conveyor, Bulletin of The Japan Society of Mechanical Engineers 1960; 3: 449-455, https://doi.org/10.1299/jsme1958.3.449.
24. Walker P, Doroszuk B, Król R. Analysis of ore flow through longitudinal belt conveyor transfer point. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (3): 536-543, https://doi.org/10.17531/ein.2020.3.17.
25. Wang S, Haolong H L, Tian R, Wang R, Xu Wang X, Sun Q, Fan J. Numerical simulation of particle flow behavior in a screw conveyor using the discrete element method, Particuology 2019; 43: 137-148, https://doi.org/10.1016/j.partic.2018.01.016.
26. Wang Y, Li T, Muzzio F J, Glasser B J. Predicting feeder performance based on material flow properties, Powder Technology 2017; 308: 135-148, https://doi.org/10.1016/j.powtec.2016.12.010.
27. White G M, Schaper L A, Ross I J, Isaacs G W. Performance characteristics of enclosed screw conveyors handling shelled corn and soybeans, Research bulletin - Purdue University Agricultural Experiment Station 1962; 740.
28. Wodołażski A. Modelling of slurry hydrodynamics with two-blade impeller in tank reactor. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(4): 533-536.
29. Zareiforoush H, Komarizadeh M H, Alizadeh M R. Performance evaluation of a 15.5 cm screw conveyor during handling process of rough rice (Oriza Sativa L.) grains, Nature and Science 2010; 8: 66-74.

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl
Article citation info:
Bukowski L, Werbińska-Wojciechowska S. Using fuzzy logic to support maintenance decisions according to Resilience-Based Maintenance concept. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 294–307, http://doi.org/10.17531/ein.2021.2.9.

Using fuzzy logic to support maintenance decisions according to Resilience-Based Maintenance concept

Indexed by: Web of Science Group

Lech Bukowski[a], Sylwia Werbińska-Wojciechowska[b,*]

[a]WSB University, ul. Zygmunta Cieplaka 1c, 41-300 Dabrowa Gornicza, Poland
[b]Faculty of Mechanical Engineering, Department of Operation and Maintenance of Technical Systems, Wroclaw University of Science and Technology, ul. Wybrzeze Wyspianskiego 27, 50-370, Wroclaw, Poland

Highlights

- Resilience-Based maintenance concept based on Maintenance Support Potentials definition.
- Implementation of a new organization's maintenance support potential level.
- New 5-grade scales for maintenance support potentials assessment.
- Fuzzy-based assessment method for organization's maintenance support capability level.

Abstract

Many authors have highlighted the importance of physical assets maintenance management in relation to resilience engineering, especially for systems operating under significant uncertainty. Thus, the authors presented a new approach to system maintenance based on resilience concept implementation. They introduced Maintenance Support Potentials (MSP) as a measure of an organization's maintenance support capacity. Moreover, based on the MSP definition, they developed a fuzzy-based organization's maintenance support potential level assessment method. The proposed approach takes into account two main MSP parameters – potential readiness level and process regency. It followed four main steps, including organization's MSP identification/evaluation, MSP weights assessment, Maintenance Support Capacity assessment, and final reasoning. A case study of a global manufacturer from the automotive industry is presented to illustrate the method's applicability. The authors also indicated further research directions to optimize the maintenance strategy based on Resilience-Based Maintenance concept.

Keywords

maintenance, system resilience, maintenance capability, fuzzy logic, uncertainty.

Abbreviations

AHP – Analytic Hierarchy Process
DE – disruptive events
FIS – fuzzy interference system
FM – forecast model
FN – fuzzy number
GC – generalized constraint
GTU – Generalized Uncertainty Theory
LP – learning process
MSC – Maintenance Support Capacity
MSP – Maintenance Support Potentials
MSS – Maintenance Support System
PI – performance indicator
PA – Potential to anticipate
PL – potential to learn
PM – potential to monitor
PR – potential to respond
RBM – Resilience-Based Maintenance
SC – selection criteria
TFN – triangular fuzzy number

1. Introduction

The maintenance of technical systems is of particular importance in the era of growing competition and ever higher requirements in quality, reliability, and productivity of organizations' functions and tasks. According to [5], maintenance for complex socio-technical systems can be defined as *a combination of activities which ensures that physical assets continue to fulfill their intended tasks effectively (performing required functions), efficiently (at minimum use of resources), and safely (at a minimum human and environmental risk)*. Therefore, the main goals of the maintenance processes of technical systems are today considered to provide [42]: 1) an appropriate level of functionality of a technical facility, 2) declared durability of a facility, 3) security of a facility and its environment, and 4) effective use of available resources supporting basic processes. The achievement of these goals is possible thanks to an appropriately selected maintenance strategy

(*) Corresponding author.
E-mail addresses: L. Bukowski - lbukowski@wsb.edu.pl, S. Werbińska-Wojciechowska - sylwia.werbinska@pwr.edu.pl

for technical systems. Currently, the primary forms of maintenance can be [25]:

- ***Pre-planned maintenance:*** includes early maintenance tasks such as cleaning, greasing, lubricating, zero-setting, and recording key measurements. Often conducted by non-maintenance staff. Usually called First-line maintenance.
- ***Planned maintenance:*** also known as scheduled maintenance, and its timing and scope are both known in advance.
- ***Shutdown maintenance:*** planned maintenance but carried out when production or plant is shut down.
- ***Breakdown maintenance:*** carried out when equipment fails to meet its desired function. This may involve repairs, replacements, or adjustments as considered necessary.
- ***Emergency maintenance*:** carried out only when either inspection or breakdown maintenance has identified its necessity.

As all technical systems operate under conditions of uncertainty and variability resulting from, among others, the uncertainty of operating processes, environment, or a modeling process, the problem of appropriate selection of the maintenance strategy arises. This is especially visible for systems performing in deep uncertainty, where the disruptive events occur very rarely. For these systems, the classical probabilistic approach to maintenance modeling cannot be implemented due to a lack of operational data.

The possible solution to this problem may be connected with the provision of resilient organizations to prevent or minimize the effects of high-level failures [23, 26].

Resilience theory is concerned with *successfully responding to the unpredictability and uncertainty of change* [4]. When referring to the resilience of industrial assets, several authors have highlighted the importance of maintenance to physical asset management and suggested ways to improve maintenance in relation to improved dependability of the assets (e.g. [29]). Moreover, the relations between maintenance, safety, risk, and resilience are especially highlighted in work [14]. Later, safety performance of organizations in relation to the decision-making processes is analyzed in [12]. The research findings constitute the base for the authors in [2], where resilience engineering issues are investigated in safety research and organizational practice. Based on the obtained survey results, in another work [20] the authors define maintenance as a safety barrier in process system operations. They analyze overall system performance in terms of, among others, maintenance costs, safety impact, environmental impact, and asset damage. Safety-II domain, defined as safety management through guided adaptability, is later investigated in [33]. The authors analyze the relations between resilience engineering and safety domains. In this new approach to safety, it is assumed that *failures were the flip side of successes, or in other words, things that go right and things that go wrong happen basically the same way*. Therefore, we may state that resilience is a key issue in ensuring the safe and reliable operation of systems and organizations' effective management.

Following this, there is a necessity to investigate the relationship between resilience and maintenance performance. For this reason, this article aims to introduce a new approach to system maintenance based on resilience concept implementation, called Resilience-Based Maintenance (RBM). The proposed concept is based on Maintenance Support Potentials (MSP) introduction. The MSP constitutes the base for measuring an organization's Maintenance Support Capacity (MSC). Moreover, based on the MSP definition, the authors develop a fuzzy-based organization's maintenance support potential level assessment method. The proposed approach considers two main maintenance support potentials parameters – potential readiness level and process regency – and four main steps, including organization's MSP and their assessment parameters identification/evaluation, MSP weights assessment, Maintenance Support Capacity in an organization assessment, and final reasoning in terms of maintenance recommendations proposition. The fuzzy theory is implemented in the MSP parameters estimation process.

The developed assessment method's implementation possibilities are based on the example of a selected global manufacturer from the automotive sector.

The proposed concept was preliminarily introduced in the authors' research work [5], where the simple investigation of maintenance potentials assessment possibilities based on scoring method was presented. In this study, the authors extend the previously done research by introducing a more systematic description of the approach, new assessment methodology, and the implementation possibilities of the RBM concept.

To sum up, the authors' contribution in this study includes:

- introduction of a new Resilience-Based Maintenance concept that bases on Maintenance Support Potentials definition,
- a new concept of organization's maintenance support potential level,
- development of a three-step assessment method to assess the organization's maintenance support potential level ratio in order to define the organization's Maintenance Support Capability,
- definition of 5-grade scales for maintenance support potentials and organization's maintenance support potential level assessment to define the maintenance support capability achieved by an organization,
- finally, the developed two-stage assessment method is implemented to verify the proposed method's diagnostic function and determine its labor intensity.

Therefore, the article structure includes, apart from the Introduction section, a detailed review of the literature in the area of classification of basic maintenance strategies for technical systems, based on which the concept of Resilience-Based maintenance is described. Next, a proposed new maintenance concept based on resilience theory is introduced. Moreover, the Maintenance Support Potentials are defined as a measure of an organization's support capability in the area of maintenance management. Later, in Section 4, the authors introduce the proposed method for an organization's maintenance support potential level assessment. The implementation possibilities of the developed method are presented in Section 5. Section 6 presents the obtained results and their discussion. Finally, Section 7 provides conclusions, limitations of the study, and suggestions for the authors' future research works to optimize the maintenance strategy based on the concept of Resilience-Based Maintenance.

2. Related work

2.1. Defining maintenance process

Today, technical systems should be designed, operated, and maintained in a safe, reliable, robust, durable, sustainable, and resilient way [34, 42]. In order to satisfy such goals, organizations develop and implement effective maintenance management processes. Following the European Standard PN-EN 13306:2010 [31], maintenance management may be defined as *all activities of the management that determine the maintenance objectives, strategies, and responsibilities and implement them by means such as maintenance planning, maintenance control and supervision, improvement of methods in the organization including economic aspects*. The main challenge for the maintenance manager is to structure maintenance procedures and activities to be undertaken to achieve the strategic objectives associated with them [6, 11]. In addition, following the European Standard EN 17007:2017 [32], *proper maintenance needs technical skills, techniques, methods to properly utilize assets like factories, power plants, vehicles, equipment, and machines* (Figure 1). As a result, it is necessary to consider maintenance issues in an organization in a more holistic way, not only limited to such problems, like maintenance planning or selection of an appropriate maintenance strategy [43]. It also requires looking at the issues related to ensuring an effective maintenance system by considering issues related to safety, risk, and resilience [19]. This is especially important for such systems, where many fluctuations due to the uncertainty may significantly influence a system's performance and its elements [18].

Therefore, the challenge of mastering uncertainty in the maintenance area seems to be the biggest problem currently facing maintenance management. To be able to do this, it is necessary to understand the nature of uncertainty and the methods for modeling it, as briefly discussed in the following Subsection 2.2.

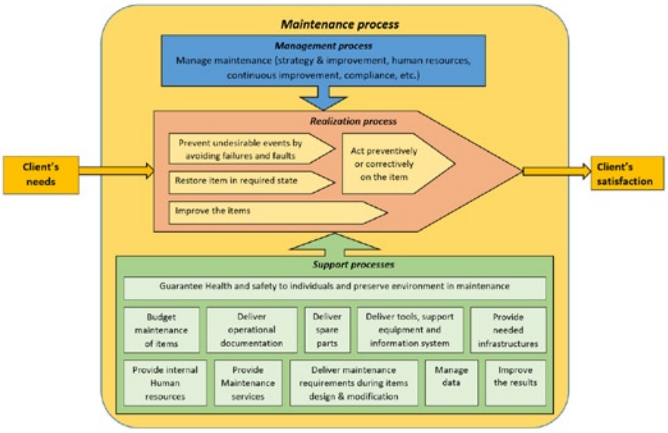


Fig. 1. Maintenance process according to the European Standard EN 17007:2017

2.2. Uncertainty modelling

The consequence of knowledge imperfections is the uncertainty in the maintenance process. We understand the concept of uncertainty as a situation of having limited knowledge such as:

- the order, nature, or state of things is unknown, and
- the consequence, extent, or magnitude of circumstances, conditions, or events is unpredictable.

There are many forms of uncertainty, but the most common is its division into two categories: aleatory and epistemic uncertainty [8, 46]. The aleatory uncertainty is understood as an inherent variation associated with the engineered system or the environment under consideration. It can be observed in random experiments and described by probability distributions. Traditional reliability engineering and risk analysis applications tend to model only the aleatory uncertainties, leading to significant underestimations of the real risks and overestimation of reliability [16]. However, the epistemic uncertainty is not an inherent property of the system or its environment, and it results from our inability to understand as well as describe and model reality. Thus, in this case, the standard probabilistic methods are not useful [41, 52].

In 2005 Lotfi A. Zadeh proposed *a generalized uncertainty theory* (GTU), which attempts to unify the approach to uncertainty [48]. The GTU theory was based on the concepts of granular structures and generalized constraints. The basic assumptions of these concepts are illustrated in Figure 2. Let X be a variable taking values in a universe of discourse, U, then a is a singular value of X (e.g., a singleton), implying that there is no uncertainty about X's value. If this is not the case, then a granular value of X, A, may be viewed as a representation of the state of knowledge about X's value.

Informally, a granule of a variable X is a clump of X values drawn together by indistinguishability, equivalence, similarity, proximity, or functionality. For example, intervals (crisp or fuzzy) are granules and different probability distributions [1]. The concept of granularity underlies the concept of a linguistic variable - a concept introduced by L. A. Zadeh in the paper "*Outline of A New Approach to the Analysis of Complex Systems and Decision Processes*" [49]. A linguistic variable's concept plays a pivotal role in many fuzzy logic applications [7, 10, 21, 30, 35, 45]. Four primary

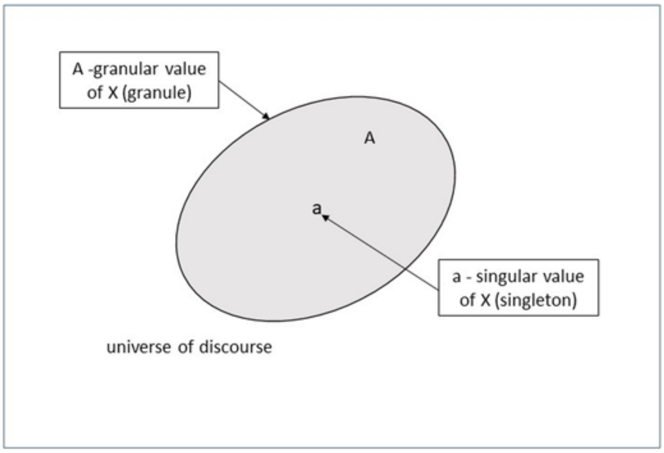


Fig. 2. Singular and granular values

rationales underlie the granulation of attributes and the use of linguistic variables:

- the bounded ability of sensory organs to resolve detail and store information,
- when numerical information is not available,
- when an attribute is not quantifiable because we do not have a numerical scale for it,
- when there is an acceptance for imperfection (e.g., inaccuracy or imprecision), which can be exploited through granulation to achieve tractability and communication economy.

There is a close connection between granularity and uncertainty. Suppose X is a variable, and we are looking for the value of this variable. If the answer is "X is a", where a is a singleton, then there is no uncertainty in X's information because the information is singular. Nevertheless, if the answer is "X is approximately a", in the abbreviation "X is $*a$", there is some uncertainty in the information because information is described as granular. Therefore, the granularity may be equated to non-singularity. In the context of standard probability theory, $*a$ would generally be interpreted as a probability distribution centered on a. In GTU, X's information is viewed as a generalized constraint on X, or more specifically, as a granule characterized by a generalized constraint. A probability distribution can be seen as a particular case of a generalized constraint.

A generalized constraint, GC, is defined as an expression of the form [47]:

$$GC : X \; isr \; R \tag{1}$$

where: X is the constrained variable; R is a constraining relation which, in general, is non-bivalent; and r is an indexing variable that identifies the modality of the constraint, that is, its semantics.

The principal modalities of generalized constraints are summarized in the following.

a) Probabilistic ($r = p$)

$$X \; isp \; R \tag{2}$$

with R – the probability distribution of X. For example:

$$X \quad isp \; N(m,\sigma \tag{3}$$

means that X is a normally distributed random variable with mean m and variance σ^2 .

If X is a random variable that takes values in a finite set $\{u_1,..., u_n\}$ with respective probabilities $p_1,..., p_n$, then X may be expressed as:

$$X \; isp\left(p_1 \backslash u_1 + \ldots + p_n \backslash u_n\right) \tag{4}$$

with the semantics:

$$Prob\left(X = u_i\right) = p_i \quad \mathrm{i = 1, \ldots , n} \tag{5}$$

In GTU, a probabilistic constraint is viewed as an instance of a generalized constraint. When X is a generalized constraint, the expression X isp R is interpreted as a probability qualification of X, with R as X's probability [47]. For example:

$$\left(X \; is \; big\right) isp \; likely \tag{6}$$

It means that the probability of the fuzzy event {X is big} is likely, where "big" is a fuzzy subset of the real line.

b) Possibilistic (r = blank)

$$X \; is \; R \tag{7}$$

with R playing the role of the possibility distribution of X. For example:

$$X \; is \left[a, b\right] \tag{8}$$

means that [a, b] is the set of possible values of X. Next example:

$$X \; is \; small \tag{9}$$

In this case, the fuzzy set labeled small is the possibility distribution of X, and μ_{small} is the membership function of small, then the semantics of "X is small" is defined by [48]:

$$Poss\left\{X = u\right\} = \mu_{small}\left(u\right) \tag{10}$$

where u is a generic value of X.

c) Veristic ($r = v$)

$$X \; isv \, R \tag{11}$$

where R plays the role of a verity (truth) distribution of X. In particular, if X takes values in a finite set $\{u_1, \ldots, u_n\}$ with respective verity (truth) values $t_1, \ldots, t_n$, then X may be expressed as:

$$X \; isv \left(t_1 | u_1 + \ldots + t_n | u_n\right. \tag{12}$$

meaning that Ver ($X = u_i$) = t_i, $i = 1, \ldots, n$.

When X is a generalized constraint, the expression X isv R is interpreted as verity (truth) qualification of X. For example:

$$\left(X \; is \; small\right) isv \; very.true \tag{13}$$

should be interpreted as „*It is very true that X is small.*" The semantics of truth qualification is defined in [47]:

$$Ver\left(X \; is \; R\right) is \; t = X \; is \; \mu_R^{-1} - 1\left(t\right) \tag{14}$$

Where μ_R^{-1} is the inverse of the membership function of R_i and t is a fuzzy truth value, which is a subset of [0, 1].

Therefore, there are two classes of fuzzy sets: (b) possibilistic and (c) veristic. In the case of a possibilistic fuzzy set, the grade of membership is the degree of possibility. In the case of a veristic fuzzy set, the grade of membership is the degree of verity (truth).

L.A. Zadeh [50] introduced the concept of fuzzy sets as a generalization of the classical set theory. In fuzzy sets, each space X element can belong partially to a set A and partly to its complement $\overline{A}$. Fuzzy sets are defined by the membership function corresponding to the functional characteristics of classical sets. Each set X element has the assigned value that defines the degree of membership to the fuzzy set. The standard fuzzy sets membership function belongs to a range [α, β] and if we deal with the normal fuzzy sets $\alpha = 0$ and $\beta = 1$. Thus, the membership function of the set X is:

$$\mu_A : X \to [0,1] \tag{15}$$

We can distinguish three cases here:

a) $\mu_A(x) = 1$ – means full membership in the fuzzy set A

b) $\mu_A(x) = 0$ – means the lack of membership in the fuzzy set A

c) $0 < \mu_A(x) < 1$ – means a partial membership in the fuzzy set A

A fuzzy set A is contained in the fuzzy set B only when $\mu_A(x) < \mu_B(x)$ for each $x \in X$, and the fuzzy set A equals the fuzzy set B only when $\mu_A(x) = \mu_B(x)$. The complement of set A is a fuzzy set $\overline{A}$ with a membership function $\mu_{\overline{A}} = 1 - \mu_A$.

Although the inference based on the fuzzy set theory and multivalued logic is more complex and less intuitive, thanks to widely available computer tools supporting the fuzzy inference process, it is becoming more common [22].

Uncertainty assessment is particularly important for *planned maintenance* and is mainly based on probabilistic models ($r = p$). However, these models' effective use is only possible if the data on the damage processes are sufficiently numerous and stationary processes. These conditions are not fulfilled in high uncertainty situations, where rare events occur, and these events' consequences are difficult to predict.

The authors propose to use possibility-based procedures to model the maintenance process under these conditions (r = blank) and fuzzy set theory. In the absence of statistical data, this approach allows objectifying expert knowledge, which is inherently subjective partially.

3. The concept of Resilience-Based Maintenance

The starting point for our considerations is the model of the maintenance process presented in Fig. 1. From this model, it follows that the prerequisite for the proper performance of the maintenance process in the organization is an extensive system implementing all support processes. We called it *Maintenance Support System (MSS)* and assumed that its fundamental characteristic is *Maintenance Support Capability (MSC)*, which is defined as follows:

Maintenance Support Capability is the ability of an organization to ensure that physical assets continue to fulfill their intended tasks effectively, efficiently, and safely, under given expected as well as unexpected conditions of use and maintenance.

Following this, a measure of an organization's Maintenance Support is its capacity to create and maintain specific potentials over time to resiliently respond to any foreseeable and unpredictable operating events.

We propose to name these potentials *Maintenance Support Potentials (MSP)* and the entire maintenance system based on this concept - Resilience-Based Maintenance. These potentials are as follows (based on [13]):

- ***PR – The Potential to respond:*** knowing what to do and being able to react correctly to any threats and hazards (e.g., changes, disturbance, and disruptions) by activating correctly planned and prepared actions, by adjusting the required mode of operation, or by introducing new activities, procedures or processes.
- ***PM – The Potential to monitor:*** being able to monitor all signals from the internal and external environment that may affect

an organization's performance in the near-term or long-term future.

- ***PL – The Potential to learn:*** being able to draw conclusions from experience, in particular '*to learn the right lessons from the right experiences*'. It also includes changing values, criteria, and even the organization's goals, depending on the type of change in the situation.
- ***PA – The Potential to anticipate:*** knowing what to be expected and predicting future developments considering particular potential disruptions, constraints, and changing operating conditions.

A functional diagram of the Maintenance Support System broken down into individual subsystems: monitoring, response, learning, and anticipation is shown in Figure 3. Thus, the general model of Maintenance Support Capability can be represented as follows:

$$MSC=\{PM,PR,PL,PA\} \tag{16}$$

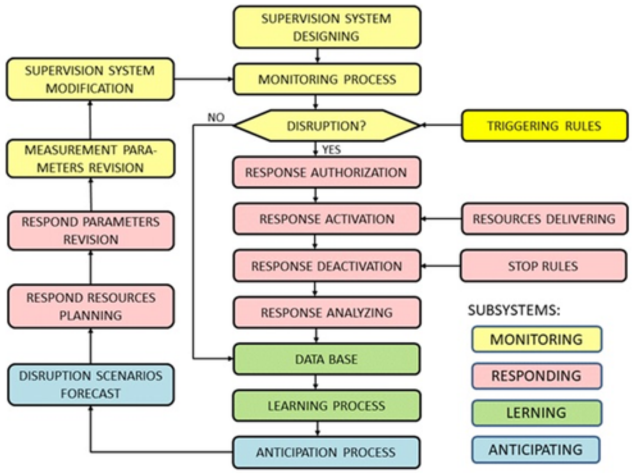


Fig. 3. A functional diagram of the Maintenance Support System

The main goal of monitoring subsystem is to improve an organization's Potential to cope with possible threats and hazards (PM). Monitoring should be proactive, recognizing upcoming situations and using information that comes from indicators that represent the current state of the performance. If the signal value from the indicator changes significantly, the response should be triggered to change the monitored system's status. The monitoring subsystem's main task is to detect disruptive situations using trigger rules and trigger a response potential (RP) when such a situation is detected. Monitoring should be carried out continuously but may change the frequency of measurements depending on the situation. In practice, a trade-off between effectiveness and accuracy of measurements is necessary. Therefore, when using the monitoring results, one must remember the uncertainty arising from this compromise.

The response to disruptive events should be both appropriate to a given situation and effective. Because no organization has infinite resources, responses can only be prepared for a limited number of disruptive events or situations. It is cost-effective to prepare a specific response for events and situations that occur frequently, but a general kind of readiness for unexpected events should be prepared. Usually, the main problem is determining the answer to two fundamental questions: *when to answer and how to answer*. Therefore, it is necessary to specify the conditions under which RBM system inputs activate (e.g., triggering rules) the response. These inputs can be seen as the outputs of a monitoring system (e.g., performance indicators).

In many cases, the timing of the Potential to respond (PR) can be critical. It is essential that the response stops neither too early nor too late. Because the triggering signal must be external to the responding subsystem, the stop rule should be internal to the response (e.g., as a part of a procedure).

Before beginning a response action, some special conditions must be fulfilled, such as requesting and receiving permission or authorization. When a response is started, the availability of specific resources should be required (e.g., information, staff, materials, and tools). While the response is being carried out, it may be necessary to maintain a given degree of normal functioning, even during an emergency action. Because responses are often complex and aggregated processes, the proper timing and synchronization of them can also be crucial in creating the Potential to respond.

Learning can be understood as the active and intended modification of processes and procedures describing the organization's behavior in specific situations. The primary purpose of Potential to learn (PL) is to improve the organization's ability to respond, monitor, and anticipate. Each organization should learn from both negative and positive examples. In general, negative situations are rare and irregular, so learning, in this case, is a reaction to some unusual event or situation (e.g., a disruption or an accident). A typical rule for starting the learning process is to state that an event or signal is significantly different from expectations. This type of learning is called reactive or event-driven.

An influential learning culture should meet four necessary conditions, namely:

- create favorable conditions and learning opportunities,
- establish main rules at which learning take place (e.g., limits and thresholds for monitored signals),
- define conditions of similarity between individual situations to enable the generalization of results obtained from monitoring,
- create objective conditions for verifying the learning process and confirming its effectiveness.

Usually, a high level of learning culture is achieved primarily by using a broad-perspective and focusing on exceptional but rare cases.

Creating Potential for anticipation (PA) in an organization is conducive to supporting anticipatory thinking technologies. Where monitoring is about observing and looking at something to see whether it is significantly changing, anticipation is more about thinking and imaging outside the event horizon. The primary purpose of anticipation is to imagine alternative scenarios and predict what can happen in the future. Therefore, anticipation depends on the assumptions made about the future and models used for prediction. Three basic types of modeling are applied in practice: deterministic, probabilistic, and realistic. The first one relies on the assumption that the future is a simple reflection of the past, both in terms of similarity in size and frequency. The basis of the probabilistic approach is the assumption that the unknown future is an extrapolation of the known past, taking into account randomness. The third method is based on the assumption that understanding past events and relationships between them makes it possible to predict the possible course of events in the future, taking into account the uncertainty that such a prediction is burdened with. Therefore, the anticipation can be seen as an art of art rather than a science and depends very much on the person or team's imagination that deals with it. This process runs at variable speeds, with unpredictable timing.

Consequently, the question arises on how to assess the maintenance support capability for an organization. To answer this question and, indeed, define the organization's maintenance support capability, there is introduced a new organization's maintenance support potential level assessment ratio (MSP_o). It can be evaluated based on the following formula:

$$MSP_o = \sum_{i=1}^{n} P_i * w_i, \quad and \quad i = 1,2,\ldots,n \tag{17}$$

where: MSP_o – organization's maintenance support potential level; P_i – ith maintenance support potential; w_i – weight for ith maintenance support potential; n – number of analyzed maintenance support potentials.

As a result, in order to gain benefit from the maintenance support potential level assessment, maintenance managers should:

1. understand MSS and identify maintenance management priorities for the near, medium, and long term;
2. identify assets, human, and material resources, as well as define the possible maintenance strategies to follow;
3. define possible responding strategies appropriate to disruptions occurring;
4. balance maintenance costs vs. risks for disruptions.

Implementation of such defined main steps for MSS identification and improvement needs a methodological approach use. Following this, the proposed approach adopted in this study consists of three main phases. The first step bases on qualitative analysis implementation. During this phase, the identification of the problem and definition of maintenance support potentials is performed. Moreover, the main parameters for maintenance support potentials assessment are identified. The second phase includes quantitative analysis performance. The collection of experts' opinions about the defined maintenance support potentials and their evaluation parameters is carried out at this stage. There are also defined weights for all maintenance support potentials to reflect a company's maintenance management priorities. Due to the lack of possibility to use accurate statistical data, it was proposed in the described method to estimate both parameters by experts using the fuzzy logic concept. The analytical approach to determine these values is presented in the 4.2 Subsection. The last phase – an output phase provides the organization's maintenance support potential level ratio assessment and reasoning on the level of MSP_o obtained.

Additionally, at this stage, the reasoning process on the maintenance recommendations proposition is performed. Figure 4 represents the graphical view of the proposed complete methodology followed.

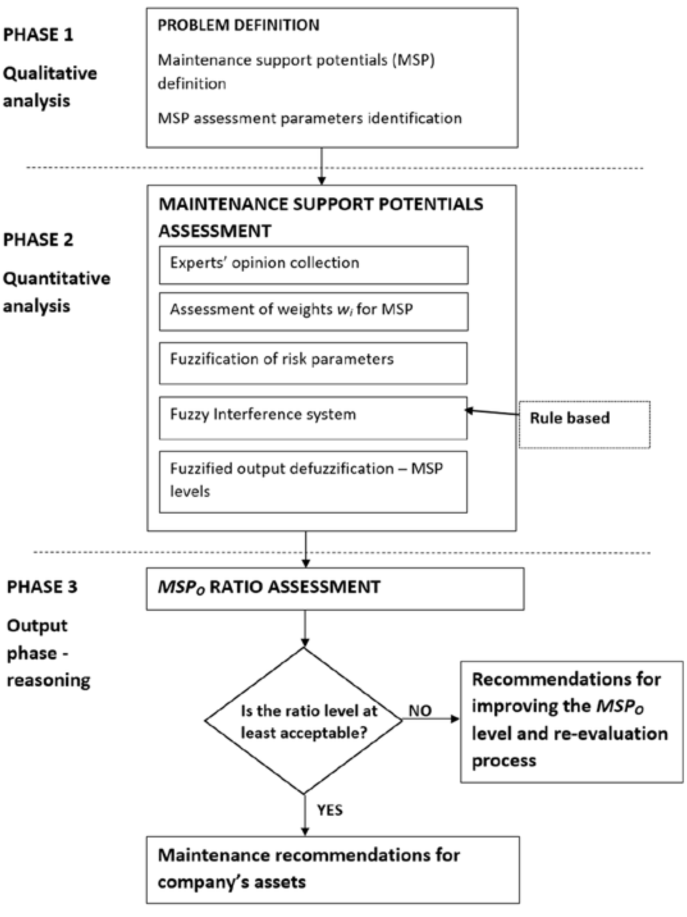


Fig. 4. Organization's maintenance support potential level analysis procedure

4. Fuzzy-based method for assessment of organization's maintenance support potential level

Before companies can devise effective means of enhancing maintenance support capacity, managers must first understand the universe of maintenance potentials as well as the conditions that drive them. Then, after gaining specific knowledge about maintenance support potentials assessment, companies can proceed to select and tailor the most effective maintenance strategies. A detailed description of the main phases of the proposed assessment method is presented in the next subsections.

4.1. Qualitative analysis of Maintenance Support Potentials

The first two steps of the analysis are used to identify the investigated maintenance support potentials and their evaluation parameters. In the developed method, the MSPs are based on resilience potentials introduced by Erik Hollnagel [13] and presented in detail in Section 3. According to [4], the MSP are usually analyzed following the six main evaluation areas. The characteristic of these areas is presented in Tables 1-4 for each MSP respectively.

Table 1. Assessment of the Potential to respond – the main evaluation areas

P_1	Procedure	Result of the procedure
p_{11}	Disruptive events (DE) identification	List of the DEs
p_{12}	Disruptive events (DE) relevance	Verified list of DEs
p_{13}	Respond to DEs planning	List of the responds to DEs
p_{14}	Respond to DEs adequacy	Verified list of the responds to DEs
p_{15}	Respond parameters defining: • triggering and ending criteria; • respond delay (activating speed); • resources capability	Verified list of the responds parameters to DEs
p_{16}	Readiness to respond	Verification rules

Table 2. Assessment of the Potential to monitor – the main evaluation areas

P_2	Procedure	Result of the procedure
p_{21}	Performance indicators (PI) identification	List of the PIs
p_{22}	Performance indicators (PI) relevance	Verified list of PIs
p_{23}	Timeliness of PIs determination	Time delay for individual PIs
p_{24}	Measurement accuracy of PIs defining	Sensitivity for individual PIs
p_{25}	Measurement frequency of PIs defining	Rules for taking measurements
p_{26}	Measurement results plausibility	Rules for checking results

The presented tables contain a detailed specification of individual factors that should be considered when assessing MSP. Such a presentation of these factors is useful for performing a preliminary analysis of the investigated organization - at the data collection stage. However, at the stage of performing a detailed quantitative analysis, such an approach would generate a very high degree of model complication. Therefore, the authors propose to group the most critical evaluation factors of individual MSP into two main parameters – potential readi-

Table 3. Assessment of the Potential to learn – the main evaluation areas

P_3	Procedure	Result of the procedure
p_{31}	Selection criteria (SC) setting	List of the SCs
p_{32}	Learning process (LP) determining	Learning process description
p_{33}	Timing of learning process determination	Time delay for implementation
p_{34}	Resources for learning process defining	Providing adequate support for LP
p_{35}	Responsibilities for LP establishing	List of responsible persons
p_{36}	Effectiveness of LP checking	Rules for checking results of LP

Table 4. Assessment of the Potential to anticipate – the main evaluation areas

P_4	Procedure	Result of the procedure
P_{41}	Forecast models (FM) elaboration	List of the FMs
P_{42}	Expertise kind and level establishing	List of the requirements
P_{43}	A time horizon of forecast determination	Time delay for individual FMs
P_{44}	Forecast accuracy defining	Uncertainty for individual FMs
P_{45}	Forecast frequency defining	Rules for taking the forecast
P_{46}	Forecast results plausibility evaluation	Rules for checking results

ness level and process regency (Fig. 5). These parameters correspond to the defined above main areas of assessment.

All the evaluation factors connected with time-frequency, timeliness or forecasting perspective refer to the regency parameter. The factors that influence organization respond capacity, measurement accuracy, learning process efficiency, or forecasting process effectiveness are attributed to the readiness parameter.

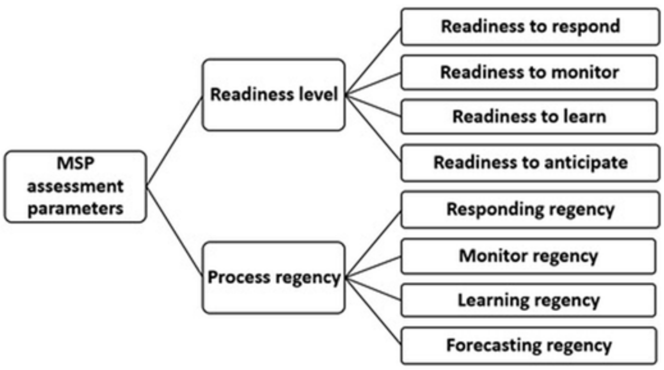


Fig. 5. MSP assessment parameters included in the proposed methodology

4.2. Quantitative analysis – assessment of Maintenance Support Potentials

During this phase, the main steps are to: collect expert opinions, weights assess, and fuzzy model implementation.

Step 1. Expert opinions collection:

First, the experts provide their opinions for the defined two MSP (P_i) assessment parameters. The experts' opinions are collected using linguistic scales. The proper definition of linguistic variables is based on expert knowledge and depends on the industry type. However, the general description of the linguistic variables is proposed in Tables 5 – 7.

Step 2. Assessment of weights for MSP:

Later, there is a necessity to assess MSP weights based on the experts' knowledge. Let $W_j = [w_1, w_2, \ldots, w_n]$ be the vector for MSP weights. Based on the available literature, this vector may be evaluated based on one of the three main approaches implementation.

Firsts approach. The parameters weights are expressed precisely by real numbers (crisp data) when satisfying the following assumption:

$$\sum_{j=1}^{n} w_j = 1 \tag{18}$$

Second approach. A vector of linguistic values may also express the weights' parameter. In this approach, there is defined the scale of linguistic terms. Thus, usually, there are used expressions to give the evaluation value of the chosen parameter by seven linguistic terms, from "Very big" to "Very small" concerning seven fuzzy scales (see, e.g. [51]). Following this, the larger weight is given to the parameter, the greater importance is given to that parameter of MSP evaluation.

Third approach. The last method of weights parameters estimation may be based on AHP method implementation. Due to the uncertainty in implementing the MSP assessment process, the fuzzy APH method should be used to find fuzzy preference weights [36]. Saaty developed the AHP method in 1980 (according to [28]). Buckley's fuzzy theory was incorporated into the AHP method in 1985 and presented in work [3]. The procedure for fuzzy AHP implementation into criteria weight evaluation is presented, e.g., in [28, 36]. According to their studies, the procedure bases on the two main steps:

- to construct fuzzy pairwise comparison matrices based on decision-makers opinion,
- to compute the fuzzy weights by normalization.

Selection of the appropriate approach for estimating the weighting factors will depend directly on the managers, their skills/expertise level, and the knowledge of possible evaluation tools. The most straightforward approach is based on the scoring method implementation but will produce a very subjective result depending on the evaluation team's preferences. The application of the AHP method will allow balancing the results obtained by assigning weights according to the level of importance of each maintenance potential in relation to the others. In turn, the second approach can be used when the assessment of the importance of individual maintenance support potentials is carried out by many experts from different departments of the company. This will unify the assessment in relation to the different levels of experience of the experts.

Step 3. Fuzzification of risk parameters:

When the expert opinions are collected and weights assessed, the fuzzy set theory is used to model the MSP parameters and obtain their assessed value. The fuzzy set theory makes the comparison process more confident [50]. Therefore, the parameters of each MSP and the output variable – MSP (P_i) level are treated as intuitionistic triangular fuzzy numbers (FN). A triangular FN is presented by a triplet $A_z = (a, b, c)$, and its membership function is given by:

$$\mu_z(x) = \begin{cases} 0 & for\ x < a \\ \dfrac{x-a}{b-a} & for\ a \le x \le b \\ \dfrac{c-x}{c-b} & for\ b \le x \le c \\ 0 & for\ x > c \end{cases} \tag{19}$$

Table 5. Process regency parameter description

Ranking category	Description
VERY HIGH (VH)	Defined and verified standards for MSP time parameters, an assessment carried out on a regular and repeatable basis.
HIGH (H)	Defined and verified standards for MSP time parameters, evaluation carried out irregularly.
MEDIUM (M)	Pre-defined and verified standards for MSP time parameters.
LOW (L)	Pre-defined standards for MSP time parameters, lack of verification processes implementation, processes are very unlikely to be evaluated in an organization.
VERY LOW (VL)	No defined standards for time parameters for MSP; an assessment may occur but will probably never be carried out.

Table 6. Readiness level description

Ranking category	Description
VERY HIGH (VH)	Fully defined and verified all processes for implementing and maintaining a given MSP in an organization.
HIGH (H)	Defined and verified procedures for MSP implementation, defined rules and principles for maintenance potential assessment without carried out a verification process.
MEDIUM (M)	Defined and verified procedures for the implementation of the Potential, lack of clearly defined rules and principles of the potential measurement procedure.
LOW (L)	Pre-defined procedures for implementing the Potential (identification of essential elements of the Potential, lack of MSP verification).
VERY LOW (VL)	Lack of defined procedures for implementing and maintaining a given potential.

Table 7. Maintenance support potential level description

Ranking category	Description
EXCELLENT (E)	Achieving and maintaining a given maintenance support potential in the organization at a very high level - readiness level and time parameters fully defined and evaluated on a regular/repeatable basis.
VERY SATISFACTORY (VS)	The parameters of a given maintenance support potential in an organization are at a high level - fully defined and evaluated on an irregular basis.
SATISFACTORY (S)	Maintenance support potential parameters at a satisfactory level - potential implementation procedures defined and verified, no rules and principles defined for evaluation, pre-defined or no standards yet being set for potential time parameters.
ACCEPTABLE (A)	Maintenance support potential parameters at an acceptable level - potential implementation procedures pre-defined, still no rules and principles established for evaluation, no standards provided for potential time parameters, probability of their evaluation pre-defined at a deficient level.
UNACCEPTABLE (UA)	Maintenance support potential parameters not defined, their evaluation nearly not possible.

The FN parameters meaning is straightforward: a and c are the lower and upper bounds of fuzzy number A_z, respectively, and b denotes the modal value of fuzzy number A_z.

If there were collected opinions from different experts, there is a necessity to aggregate them to obtain the P_i level. According to [9], the aggregation of exerts opinion can be performed using the arithmetic mean aggregation operator. The mean aggregation operator, defined on fuzzy triangular numbers (a_1, b_1, c_1), (a_2, b_2, c_2)… (a_m, b_m, c_m), delivers the result as (x, y, z) according to the formula:

$$\begin{cases} x = \dfrac{1}{m}\sum\limits_{k=0}^{m} a_k \\ y = \dfrac{1}{m}\sum\limits_{k=0}^{m} b_k \\ z = \dfrac{1}{m}\sum\limits_{k=0}^{m} c_k \end{cases} \tag{20}$$

For the transparency of the presented method, the authors do not consider experts weighting. However, when there is a need to differentiate the obtained opinions depending on an expert's significance, the authors recommend introducing the experts' normalized weights. Thus, the aggregated fuzzy number of the ith basic opinion may be estimated as [27]:

$$M_{zi} = \sum_{l=1}^{m} A_{zl}, and\ i = 1,2,\ldots,n_p \tag{21}$$

where: M_{zi} represents aggregated fuzzy number of the ith parameter; W_l is an lth experts' normalized weight; A_{zil} is the fuzzy number of ith parameter given by lth expert judgment; n_p is the number of parameters; m is the number of experts.

According to the expert's trait, the examples of weighting scores are presented, e.g., in [27, 44]. Moreover, the survey of known methods for fuzzy opinions aggregation is given in, e.g. [15, 39].

Step 4. Fuzzy interference system:

After carrying out experts' opinions aggregation, the next steps of this phase are connected with P_i quantification. This process is based on a Mamdani fuzzy model use [24]. The Mamdani fuzzy interference mechanism is based on the compositional rule of inference proposed by Zadeh [50]. The Mamdani fuzzy model's main components are Fuzzification, Knowledge base, Fuzzy Interference System, and Defuzzification [38]. A scheme of the P_i assessment process based on the Mamdani fuzzy model is shown in Figure 6.

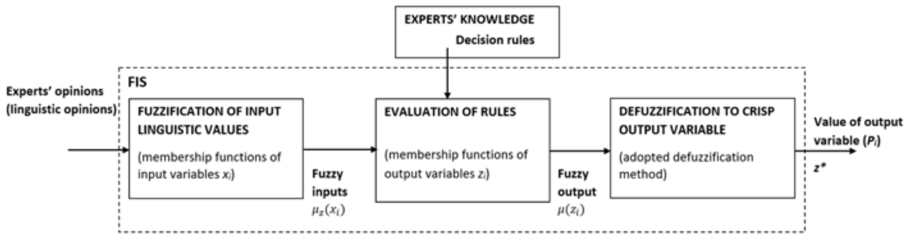


Fig. 6. Stages of the use of fuzzy sets according to Mamdani fuzzy model

As it was mentioned, the fuzzification process is based on the use of TFN. A triangular FN converts the linguistic scales in the range of 0-1 using its membership function. The knowledgebase consists of the rule base and membership functions of inputs. The rule base includes a number of *IF-THEN* rules used to capture the imprecise modes of reasoning [40].

The fuzzy interference system (FIS) is designed to map the fuzzy inputs and rule the outputs using a fuzzy set theory. Due to the Mamdani model use, the FIS has based on MIN and MAX operators implementation. The MIN operator is used for combination and implication operations. The MAX operator is used to aggregate the fuzzy outputs.

Finally, the defuzzification process is aimed at the conversion of the fuzzy output into crisp output.

Step 5. Fuzzified output defuzzification:

A survey of the most commonly known defuzzification methods is presented, e.g., in [37]. There are many sources of uncertainty in evaluating MSP parameters, so the authors propose using the centroid of area defuzzification method for defuzzification process performance. Thus, the crisp output is estimated as [38]:

$$Centroid\ of\ area, z^* = \frac{\int \mu_A(z) \cdot zdz}{\int \mu_A(z)dz} \tag{22}$$

where: z^* – the crisp value for the z output (defuzzified output); $\mu_A(z)$ – the aggregated output membership function; z – universe of discourse.

This crisp output value is later implemented in the Output phase for MSP_o level estimation.

4.3. Organization's maintenance support potential level assessment with the reasoning process

The organization's maintenance support potential level is estimated based on the previous phases' results in the last phase. According to the obtained MSP_o level, maintenance-related decisions can be made accordingly.

Based on the obtained level of estimated ratio, the decision-maker should firstly correctly interpret the obtained values. Table 8 describes the possible MSP_o levels. The authors propose a 5-grade scale for MSP_o ratio assessment.

According to the obtained overall ratio level, the decision-maker may take appropriate actions. When the overall ratio level is **not acceptable**, managers should take the following actions:

- first, the definition of maintenance management policies and procedures as a basis for MSP implementation,
- introduction of disruptive events identification processes and possible responding parameters definition; establishing the possible influence of adverse events occurrence on maintenance processes performed in an organization,
- maintenance measurement processes definition with a selection of possible performance indicators,
- analysis of possible to be implemented in organization forecast models, which provide the most efficient maintenance management process.

This means the manager must seek additional management actions for company maintenance support capability introduction and improvement or increase prevention and preparedness (connected with, e.g., maintenance policy definition) without reducing profits. Success at this task requires a good understanding of organization's maintenance support system, both broad and tailored to the manager's own company. Moreover, it constitutes the initial step of MSS creation in an organization.

For organizations where the overall ratio assumes values accepted by managers, decisions concerns maintenance recommendations. When the overall ratio is acceptable, the most common maintenance recommendations are the following maintenance policies defined by a producer. When the obtained level of an overall ratio is higher than the acceptable level, the organization maintenance capability is enough to introduce maintenance strategies that satisfy reliability or risk/ safety assumptions. The appropriate recommendations will depend on the type of organization, its physical assets, and industry sector and should be compatible with ISO 5500x standards indications.

Following this, the general structure for MSC in organization development may be compatible with the one presented in Fig. 7. The most crucial improvement ways are indicated in every of the analyzed maintenance support potentials.

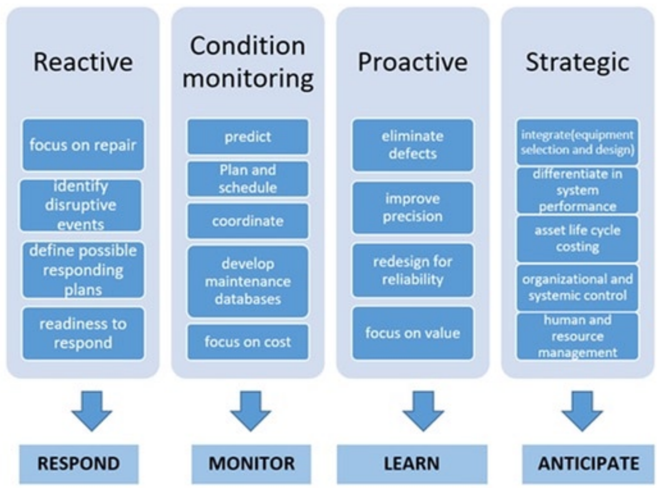


Fig. 7. MSC in organization development – possible directions of company's related tasks

5. Application of the proposed approach in a company from the automotive sector

To illustrate the proposed fuzzy-based decision method's implementation possibility, the authors analyzed a case company from the automotive sector. The investigated company is located in Poland in the Lower Silesia region and is a global manufacturer of compres-

Table 8. *Organization's maintenance support potential levels – description and ratio levels*

Ranking category	Description	MSP_o range
EXCELLENT (E)	Full development and implementation of maintenance support potentials in the organization; a system for collecting and using information (about adverse events as well as processes for responding to their occurrence) following the concept of a learning organization. Integration of a maintenance management system with an enterprise management strategy. Parameters of maintenance support potentials in an organization evaluated regularly.	93-100
VERY SATISFACTORY (VS)	The parameters of a given maintenance support potential in an organization are fully defined and implemented, the monitoring of the level of maintenance support potentials is based on a defined system of operational indicators, the evaluation process is still carried out on an irregular basis; however, information on potential adverse events is collected in a systematic manner.	75-92
SATISFACTORY (S)	Maintenance support potential parameters at a satisfactory level - potential implementation procedures are defined and verified, there still are no rules and principles for potentials evaluation, a system for measuring maintenance support potentials is still not fully developed; standards for potential time parameters are not defined or just pre-defined.	54-74
ACCEPTABLE (A)	Maintenance support potential parameters at an acceptable level - procedures for implementation of potential are pre-defined; there are no rules and principles for evaluation of maintenance potentials, but possible undesirable events are preliminarily identified; there are no standards for time parameters of Potential, and a probability of their evaluation is estimated at a deficient level.	31-53
UNACCEPTABLE (UA)	No activities are carried out to implement and evaluate maintenance support potentials in an organization; no efforts (or very little) are made to identify adverse events and their impact on maintenance processes, no management policies and procedures in the maintenance area.	0-30

Table 9. *Parameters linguistic scores for all defined MSP based on experts' opinions*

MSP	Process regency level	Readiness level
P_1 – Potential to respond	VH	H
P_2 – Potential to monitor	H	H
P_3 – Potential to learn	M	H
P_4 – Potential to anticipate	M	M

sors for automotive air conditioning. The company was launched in Poland in 2005. Each year, it produces about 3 million compressors delivered to European assembly plants of the biggest car manufacturers of world-famous brands such as Volkswagen, Volvo, or Ford. The analyzed production plant currently operates 26 production lines, including processes such as high-precision machining, grinding, electron welding, friction welding, and coating.

The company's primary goal is to respond to the customers' demand for appropriate technologies, products, and services. World's success is based on the three strategic pillars: quality, cost and delivery on time, and continuous product development with constant care for the environment. The achievement of these policy goals is connected with conducted some priority actions in the company. One of them is connected with risk-based thinking and continuous improvement of an integrated management system. The main goals of the risk management system adopted in the company is to ensure proper performance of its goals and tasks and create the company's resilience system. Additionally, the risk is defined as the effect of uncertainty on objectives. The currently implemented risk management system focuses on 13 main areas (e.g., Business risk management, Legal risk management, Occupational risk management, Environmental risk management, Operational (production/logistic) risk management, and Supply risk management). The adopted company's approach to risk management is structured and is compatible with ISO 31000 standard [17]. Analyzes are carried out on an ongoing basis, and the results are continuously monitored. The introduced risk management approach is based on the simplified FMEA (Failure Mode and Effect Analysis) method.

Following this, the authors analyze if the company, which is focused on risk management and safety issues, follows the main resilience potentials according to the RBM concept. This gives the possibility to make a statement of the new resilience engineering-based approach implementation possibilities. The evaluation of the analyzed organization's maintenance support capability level was conducted using the fuzzy rule-based risk assessment method presented in Section 4. Moreover, the developed assessment method's implementation process was carried out using the fuzzy logic toolbox of MATLAB version R2020a. The main implementation phases of the assessment method are presented below.

First, the quantitative analysis was performed. The surveyed company's experts gave their opinions. The obtained MSP parameters linguistic scores are presented in Table 9.

Moreover, it was assumed that all MSPs have the same importance for organization maintenance support capacity level assessment. Following this, the weights of the parameters are expressed precisely by real numbers (crisp data), and all are equal to wi = 0.25 (according to Equation (18)).

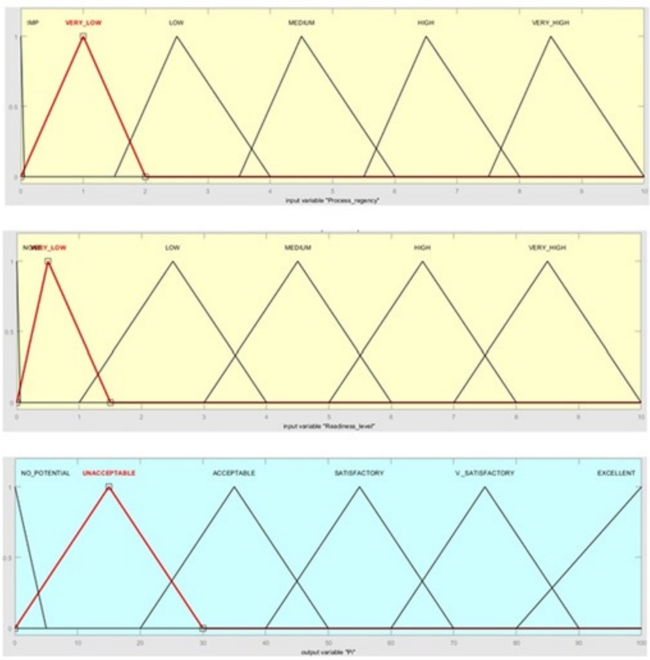


Fig. 8. *Membership functions of a) process regency, b) readiness level, and c) MSP (P_i) level*

In the next step, the proposed fuzzy model needs to be implemented. Following this, the input parameters are to be fuzzified. The obtained linguistic scores given by the experts are converted to corresponding fuzzy set numbers. The Triangular and Trapezoidal FNs used in the presented case study to represent the linguistic scales of input and output parameters are shown in Figure 8.

Next, there is a necessity to determine IF-THEN rules. Based on the experts' knowledge, there were proposed 25 rules – presented in Table 10 and one additional, which defines the situation when there is no potential identified (rule 26). According to this, for example, rule 1 is defined as:

IF Process regency is Very Low and Readiness level is Very Low, THEN P_i level is Unacceptable.

Rule 26 is defined as:

If Process regency is Impossible and Readiness level is None, THEN P_i level is NO POTENTIAL

Table 10. MSP level decision matrix

Process regency	VH	S	S	VS	E	E
	H	A	S	VS	VS	E
	M	A	A	S	S	VS
	L	UA	A	A	S	S
	VL	UA	UA	A	S	S
		VL	L	M	H	VH
		Readiness level				

With MATLAB software, there is possible to obtain the final MSP score from the constructed FIS. Figure 9 presents the adopted rules in the used MATLAB software for chosen P_i assessment. To obtain Pi's final score from the constructed FIS, Equation (22) is used for the defuzzification of the fuzzy set resulting from the Mamdani algorithm.

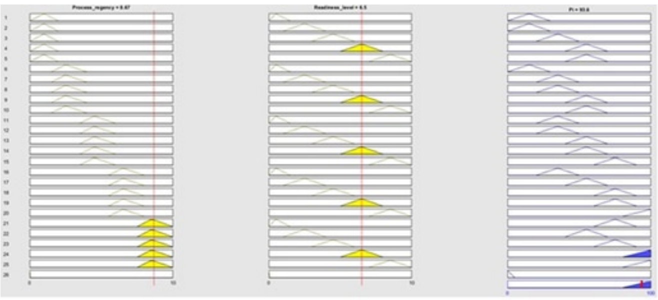


Fig. 9. Sample rule base for maintenance support potential P_1 assessment

Table 11 presents the obtained organization's maintenance support potential level for the analyzed company.

Table 11. Evaluated organization's maintenance support potential level

	P_i	$P_i{*}w_i$
P_1 – Potential to respond	91.36	22.84
P_2 – Potential to monitor	75	18.75
P_3 – Potential to learn	55	13.75
P_4 – Potential to anticipate	55	13.75
MSP_o		69.09

The results given in the Table 11 are obtained considering the assumptions that all decision rules have the same weights (the same importance). The discussion of obtained results is presented in the next section.

6. Results and discussion

The proposed case study gives the possibility to analyze how the developed fuzzy-based assessment method may be used to evaluate the maintenance support capability level in an organization. The proposed method allows the possibility to employ linguistically exert knowledge and engineering judgment to make a more realistic evaluation in the maintenance management capability area. The complete results of the proposed FIS for MSP assessment are presented in Figure 10.

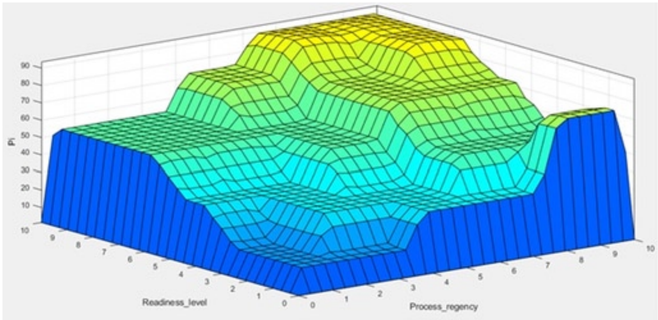


Fig. 10. Surface view of the proposed fuzzy inference system (all rules are with weight = 1)

This 3D plot shows the resultant values of preliminarily estimated MSP parameters – readiness level and process regency. The readiness level can be understood as an ability to maintain a system with the necessary resources, the possibility of reacting for occurred failure, and no reputational damage occurrence. The process regency corresponds to the regularity of performed monitoring and learning processes connected with maintenance management performance in an organization. The lowest, dark blue part of the plot represents the resultant low level of MSP resulting from low readiness levels or lack of regularity of process performance, allowing for disruptive events identification, processes monitoring, or forecasting.

It should be mentioned that the uppermost corner, the yellow field, represents theoretically the highest score – the excellent level of maintenance support potential in an organization (10/10 on the probability scale), which would provide the highest maintenance capability in organization achievement.

Following this, the plot provides a quantified basis for making managerial decisions regarding taking up active methods to improve the MSP level or observing its level to determine when such actions should be implemented. Therefore, a management plan can be prepared accordingly so that preventive actions can be taken up for the riskiest/the most disruptive events. As a result, the safety of an organization may be improved.

According to the presented results for the analyzed production company, the obtained level of overall ratio in the organization is about 70; hence, based on the description given in the Table 8, the obtained organization's maintenance support potential level is satisfactory. The assessment coincides with the authors' observations, wherein the company has developed a risk management system, but the results are not translated into decisions in the area of technical maintenance.

The first MSP – Potential to respond has obtained the highest level during the evaluation process. This is mainly connected with a well-developed risk management system that clearly identifies potential internal and external risks. In addition, respond parameters and response plans have been defined.

The second maintenance support potential – Potential to monitor has also been highly rated. The analyzed company monitors risks/opportunities in an ongoing manner and follows business continuity. Moreover, it has an extensive performance measurement system, especially in the area of production management. The primary measures allow for precise identification of disruptive events and their parameters/consequences (e.g., duration/removal from the system, costs, and delays). The frequency of PI's defining and monitoring is also defined.

Despite implementation in the case company of systematic, periodic analysis of current business performance and ways to address risks, the Potential to learn still needs to be improved. This company's maintenance support potential still needs to be supplemented with, among others, possible resources for the learning process and its time parameters.

Moreover, according to the expert opinions, the case company's Potential to anticipate is at the medium level. This is mainly due to its focusing on the current state of the operational performance level. There is a lack of solutions focused on predicting future developments on particular potential disruptions, constraints, and changing operating conditions. The leading implemented solutions assess the analyses of the company's current operational/strategical and tactical level made once a year (in the last month of the fiscal year).

Following this, according to the obtained results, there is still the necessity to define processes that would predict potential future adverse events with a significant impact on the implemented operational processes. There is a lack of solutions that would focus on predicting future development direction in relation to specific potential disruptions, constraints, and changing operating conditions.

Consequently, the main recommendations for the analyzed company regarding its maintenance processes are as follows:

- use the results obtained from risk management in the planning of effective maintenance processes (maintenance strategies) to help achieve the required availability, reliability, and safety levels dictated by the business,
- develop guidelines for the forecasting process, preventive procedures, and maintenance management scenarios for identified disruptive events,
- assess strategy elements aimed at achieving the required availability, reliability, and safety levels dictated by the business (e.g. critical spare management, operational controls, and failure response measures),
- provide transparently and verifiably costing in the area of maintenance management,
- embrace and develop approaches that seek to continually improve efficiency and effectiveness of company's activities (e.g. connected with learning objectives description, learning process rules definition),
- be compliant with statutory and regulatory imperatives.

7. Conclusions

Maintenance management is one of the most important issues nowadays. The appropriate maintenance decisions can achieve significant financial benefits (reducing maintenance costs) and increasing the company's operational indicators. However, when defining an organization's maintenance support capability, the focus should be on eliminating possible hazards and preventing failures, and developing an organization's potentials for resilient performance. Following this, the authors introduced a new concept on Resilience-Based Maintenance, which is based on implementing the main resilience potentials given by Erik Hollnagel [13]. Its proper use in a company is based on the necessity of its performance investigation in such areas as current state and knowledge about possible hazard events, the possibility of learning from the obtained experience, or the ability to anticipate unwanted events. Moreover, in this study, the authors proposed the organization's maintenance support potential level assessment methodology, which considers two evaluation parameters – readiness level and process regency. The fuzzy logic structure allowed the experts to capture the experts' opinions in linguistic terms for the defined two MSP parameters and evaluate the overall ratio level.

The analyzed case study shows the possibilities of using a given method in the decision-making process. Hence, the selected case's assessment procedure allowed us to verify the complexity of the adopted procedure, the substantive scope of the developed assessment tool and allowed us to determine the intensity of implementation work.

Following the case study, we may state that the analyzed company is well prepared to respond to everyday hazards. The main problem is connected with developing such tools and skills that will give the possibility to predict the future.

At this stage of carried out research analyses, the authors may point out two main possible limitations of the proposed method. First, the method limitation may be connected with the managers' correctness of performed assessment process. The managers (experts) may give incorrect answers during the internal audit performance to obtain higher ratings than the actual level of achieved maintenance management capability in an organization. The second possible limitation is the possibility of omission of specific steps during the proposed assessment procedure performance. Following this, to obtain reliable results, it is necessary to follow the procedure and appropriately evaluate the actual level of the maintenance support potentials being assessed in the model.

The results presented in the article are preliminary studies that the authors will develop in their future research. Further analysis will focus on the business continuity concept implementation and physical asset management concept use to extend the proposed Resilience-Based Maintenance approach.

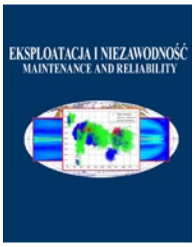
To conclude, the proposed methodology is to be used for organization maintenance support capability level assessment and may be performed by maintenance management and safety officers. Moreover, it gives preliminary information that can be useful for the development of maintenance strategies as well as the selection of the most hazardous areas in the audited companies. Therefore, it provides essential information on the need to control disruptive events and implement safety improvements. Moreover, the proposed organization's maintenance support potential level assessment method may be used in various industry sectors.

References

1. Bargiela A, Pedrycz W. Granular Computing. Kluwer Academic Publishers: 2002, https://doi.org/10.1007/978-1-4615-1033-8.
2. Bergstrom J, van Winsen R, Henriqson E. On the rationale of resilience in the domain of safety: a literature review. Reliability Engineering and System Safety 2015; 141: 131-141, https://doi.org/10.1016/j.ress.2015.03.008.
3. Buckley JJ. Fuzzy hierarchical analysis. Fuzzy Sets and Systems 1985; 17(3): 233-247, https://doi.org/10.1016/0165-0114(85)90090-9.
4. Bukowski L. Reliable, Secure and Resilient Logistics Networks. Delivering products in a risky environment. Springer Nature Switzerland AG: 2019, https://doi.org/10.1007/978-3-030-00850-5.
5. Bukowski L, Werbińska-Wojciechowska S. Resilience based maintenance: a conceptual approach. In: Baraldi P, Di Maio F, Zio E. (eds): Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference, Research Publishing, Singapore: 2020: 3782-3789. https://doi.org/10.3850/978-981-14-8593-0.
6. De Almeida AT, Cavalcante CAV, Alencar MH, Ferreira RJP, De Almeida-Filho AT, Garcez TV. Multicriteria and multiobjective models for risk, reliability and maintenance decision analysis. Springer International Publishing Switzerland: 2015, https://doi.org/10.1007/978-3-319-17969-8_12.
7. Dubois D, Prade H. (Eds.) Fuzzy Information Engineering: a Guided Tour of Applications. John Wiley and Sons: 1996.

8. Eusgeld I, Freiling F C. Introduction to Dependability Metrics. In: Eusgeld I, Freiling F C, Reussner R. (eds.) Dependability Metrics. Lecture Notes in Computer Science, vol 4909. Springer, Berlin, Heidelberg: 2008: 1-4, https://doi.org/10.1007/978-3-540-68947-8_1.
9. Fasanghari M, Roudsari FH. The fuzzy evaluation of e-commerce customer satisfaction. World Applied Sciences Journal 2008; 4(2): 164-168, https://doi.org/10.1109/ISECS.2008.207.
10. Filev D, Yager RR. Essentials of Fuzzy Modeling and Control. Wiley-Interscience: 1994.
11. Gandhare BS, Akarte M. Maintenance strategy selection. In: Proc. of Ninth AIMS International Conference on Management, January 1-4, 2012: 1330-1336.
12. Hayes J. Use of safety barriers in operational safety decision making. Safety Science 2012; 50: 424-432, https://doi.org/10.1016/j.ssci.2011.10.002.
13. Hollnagel E. Safety-II in Practice. Developing the Resilience Potentials. Routledge, Taylor & Francis Group, London and New York: 2018, https://doi.org/10.4324/9781315201023.
14. Hollnagel E. Risk + barriers - safety? Safety Science 2008; 46: 221-229, https://doi.org/10.1016/j.ssci.2007.06.028.
15. Hsu H-S, Chen Ch-T. Aggregation of fuzzy opinions under group decision making. Fuzzy Sets and Systems 1996; 79(3): 279-285, https://doi.org/10.1016/0165-0114(95)00185-9.
16. Huang H Z. Structural reliability analysis using fuzzy sets theory. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14 (4): 284-294.
17. ISO 31000:2018 Risk management - Guidelines. International Organization for Standardization, Geneva.
18. Jacyna M, Semenov I. Models of vehicle service system supply under information uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (4): 694-704, https://doi.org/10.17531/ein.2020.4.13.
19. Jain P, Mentzer R, Mannan MS. Resilience metrics for improved process-risk decision making: Survey, analysis and application. Safety Science 2018; 108: 13-28, https://doi.org/10.1016/j.ssci.2018.04.012.
20. Jain P, Pistikopoulos EN, Mannan MS. Process resilience analysis based on data-driven maintenance optimization: Application to cooling tower operations. Computers and Chemical Engineering 2019; 121: 27-45, https://doi.org/10.1016/j.compchemeng.2018.10.019.
21. Jamshidi M, Titli A, Zadeh LA, Boverie S. (Eds.) Applications of Fuzzy Logic-towards High Machine Intelligence Quotient Systems. Environmental and Intelligent Manufacturing Systems Series, vol. 9, Prentice Hall, Upper Saddle River, NJ: 1997.
22. Klir G J, Yuan B. Fuzzy sets and fuzzy logic: theory and applications. Upper Saddle River, NJ: Prentice Hall, PTR: 1995.
23. Loska A. Remarks about modelling of maintenance processes with the use of scenario techniques. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14 (2): 92-98.
24. Mamdani EH, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies 1975; 7: 1-13, https://doi.org/10.1016/S0020-7373(75)80002-2.
25. Misra KB. Maintenance Engineering and Maintainability: An Introduction. In: Misra KB (eds): Handbook of Performability Engineering. Springer, London: 2008: 755-772, https://doi.org/10.1007/978-1-84800-131-2_46.
26. Moerman J-J, Braaksma J, van Dongen L. Resilient performance in maintenance operations: managing unexpected failures. In EurOMA 2017 conference proceedings. 24th International Annual EurOMA Conference 2017, Edinburgh, United Kingdom, 1/07/17, 1-10.
27. Mottahedi A, Ataei M. Fuzzy fault tree analysis for coal burst occurrence probability in underground coal mining. Tunnelling and Underground Space Technology 2019; 83: 165-174, https://doi.org/10.1016/j.tust.2018.09.029.
28. Nadaban S, Dzitac S, Dzitac I. Fuzzy TOPSIS: A general view. Proceedia Computer Science 2016; 91: 823-831, https://doi.org/10.1016/j.procs.2016.07.088.
29. Okoh P, Haugen S. Improving the robustness and resilience properties of maintenance. Process Safety and Environmental Protection 2015; 94: 212-226, https://doi.org/10.1016/j.psep.2014.06.014.
30. Pedrycz W, Gomide F. Introduction to Fuzzy Sets. MIT Press, Cambridge, MA: 1998, https://doi.org/10.7551/mitpress/3926.001.0001.
31. PN-EN 13306:2010 Maintenance - Maintenance terminology. European Committee for Standardization, Bruxelles.
32. PN-EN 17007:2018-02 Maintenance process and associated indicators. European Committee for Standardization, Bruxelles.
33. Provan DJ, Woods DD, Dekker SWA, Rae AJ. Safety II professionals: How resilience engineering can transform safety practice. Reliability Engineering and System Safety 2020; 195: 1-14, https://doi.org/10.1016/j.ress.2019.106740.
34. Rau C-G, Necas P, Boscoianu M. Review of maintainability and maintenance optimization methods for aviation engineering systems. Science and Military 2011; 2: 54-60.
35. Ross T J. Fuzzy Logic with Engineering Applications. John Wiley and Sons: 2004.
36. Sun Ch-Ch. A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS methods. Expert Systems with Applications 2010; 37: 7745-7754, https://doi.org/10.1016/j.eswa.2010.04.066.
37. Talon A, Curt C. Selection of appropriate defuzzification methods: application to the assessment of dam performance. Expert Systems with Applications 2017; 70: 160-174, https://doi.org/10.1016/j.eswa.2016.09.004.
38. Tripathy DP, Ala CK. Risk assessment in underground coal mines using fuzzy logic in the presence of uncertainty. Journal of The Institution of Engineers (India): Series D 2018; 99(1): 157-163, https://doi.org/10.1007/s40033-018-0154-7.
39. Vaníček J, Vrana I, Aly S. Fuzzy aggregation and averaging for group decision making: A generalization and survey. Knowledge Based Systems 2009; 22: 79-84, https://doi.org/10.1016/j.knosys.2008.07.002.
40. Verma S, Chaudhri S. Integration of fuzzy reasoning approach (FRA) and fuzzy analytic hierarchy process (FAHP) for risk assessment in mining industry. Journal of Industrial Engineering and Management 2014; 7: 1347-1367, https://doi.org/10.3926/jiem.948.
41. Wang H, Duan F, Ma J. Reliability analysis of complex uncertainty multi-state system based on Bayesian network. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (3): 419-429, https://doi.org/10.17531/ein.2019.3.8.
42. Werbińska-Wojciechowska S. Technical system maintenance. Delay-Time-Based Modelling. Springer Nature Switzerland AG: 2019, https://doi.org/10.1007/978-3-030-10788-8.
43. Yan S, MA B, Wang X, Chen J, Zheng C. Maintenance policy for oil-lubricated systems with oil analysis data. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (3): 455-464, https://doi.org/10.17531/ein.2020.3.8.
44. Yazdi M, Nikfar F, Nasrabadi M. Failure probability analysis by employing fuzzy fault tree analysis. International Journal of System Assurance Engineering and Management 2017; 8: 1177-1193, https://doi.org/10.1007/s13198-017-0583-y.
45. Yen J, Langari R. Fuzzy Logic: Intelligence, Control and Information. Prentice Hall, Berlin: 1998.

46. You L, Zhang J, Li Q, Ye N. Structural reliability analysis based on fuzzy random uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (4): 599-609, https://doi.org/10.17531/ein.2019.4.9.
47. Zadeh L A. Toward a generalized theory of uncertainty (GTU) - an outline. Information Sciences 2005; 172: 1-40, https://doi.org/10.1016/j.ins.2005.01.017.
48. Zadeh L A. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Systems 1978; 1: 3-28, https://doi.org/10.1016/S0165-0114(99)80004-9.
49. Zadeh L A. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man and Cybernetics SMC-3; 1973: 28-44, https://doi.org/10.1109/TSMC.1973.5408575.
50. Zadeh L A. Fuzzy Sets. Information and Control 1965; 8: 338-353, https://doi.org/10.1016/S0019-9958(65)90241-X.
51. Zheng J H. A fuzzy TOPSIS approach based to evaluate the transportation mode selection: an experience in a suburban university. Advances in Transportation Studies, Special Issue 2015; 1: 23-34, https://doi.org/10.4399/978885488881403.
52. Zhiyong G, Jiwu L, Rongxi W. Prognostics uncertainty reduction by right-time prediction of remaining useful life based on hidden Markov model and proportional hazard model. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23 (1): 154-164, https://doi.org/10.17531/ein.2021.1.16.



Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl
Article citation info:
Huang T, Xiahou T, Li Y-F, Qian H-M, Liu Y, Huang H-Z. Reliability assessment of wind turbine generators by fuzzy universal generating function. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 308–314, http://doi.org/10.17531/ein.2021.2.10.

Reliability assessment of wind turbine generators by fuzzy universal generating function

Indexed by: Web of Science Group

Tudi Huang[a], Tangfan Xiahou[a], Yan-Feng Li[a], Hua-Ming Qian[a], Yu Liu[a], Hong-Zhong Huang[a,*]

[a]Center for System Reliability and Safety, School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China Sichuan, 611731, P. R. China

Highlights

- The fuzzy states of the DFIG systems are provided.
- All components' states are given as triangular fuzzy number based on experts' experience.
- The reliability assessment of the DFIG based on the FUGF is performed.

Abstract

Wind power has been widely used in the past decade because of its safety and cleanness. Double fed induction generator (DFIG), as one of the most popular wind turbine generators, suffers from degradation. Therefore, reliability assessment for this type of generator is of great significance. The DFIG can be characterized as a multi-state system (MSS) whose components have more than two states. However, due to the limited data and/or vague judgments from experts, it is difficult to obtain the accurate values of the states and thus it inevitably contains epistemic uncertainty. In this paper, the fuzzy universal generating function (FUGF) method is utilized to conduct the reliability assessment of the DFIG by describing the states using fuzzy numbers. First, the fuzzy states of the DFIG system's components are defined and the entire system state is calculated based the system structure function. Second, all components' states are determined as triangular fuzzy numbers (TFN) according to experts' experiences. Finally, the reliability assessment of the DFIG based on the FUGF is conducted.

Keywords

reliability assessment, double fed induction generator, multi-state system, fuzzy universal generating function.

Acronyms and Abbreviations

DFIG Double Fed Induction Generator
MSS Multi-state System
UGF Universal Generating Function
FUGF Fuzzy Universal Generating Function
TFN Triangular Fuzzy Number
PMG Permanent Magnet Generator
PD Probability Distribution

1. Introduction

Energy is closely connected with our human beings. With the increasing crisis on energy and environmental problems, wind energy has gained significant attention in recent years due to its safety and cleanness. Consequently, technologies related to wind energy developed fast in the past decade. With the increasing capacity of wind turbine generators, wind turbine generator systems are becoming more and more compounded and complicated, especially for the megawatt-scale wind turbine generator systems. For a large complex equipment such as wind turbine generator systems, much attention should be paid to their reliability assessment besides considering their capacity. In general, the designed life span of a wind turbine generator is 20 years. Thus, it is very difficult to have an accurate reliability assessment of the wind turbine.

Due to the external working environments and internal failure dependence, the double fed induction generator (DFIG), as a typical type of wind turbines, inevitably deteriorates with the usage. Once the deterioration beyond the acceptable level, it is deemed as failure. The failure of the DFIG will not only cause energy loss but also create damage to the entire wind farm. Therefore, reliability assessment for the DFIG is of great significance. In the literature, many works on reliability assessment of the DFIG have been reported. Carroll et al [1] studied the reliability of wind turbines with the DFIG and permanent magnet generator (PMG) drive trains. Zhou et al. [34] conducted certain attempts on reliability and performance improvement of the DFIG. Note that most of the existing studies assumed the DFIG system and its components as a binary-state system or components, i.e., the working state and the failure state. However, the DFIG is typically made up of five main components, i.e., blades, gearboxes, generators, converters and transformers. Blades and gearboxes are mechani-

(*) Corresponding author.
E-mail addresses: T. Huang - huangtudi@std.uestc.edu.cn, T. Xiahou xiahoutf@163.com, Y-F. Li - yanfengli@uestc.edu.cn, H-M quian - qianhuaming@std.uestc.edu.cn, Y. Liu - yuliu@uestc.edu.cn, H-Z Huang - hzhuang@uestc.edu.cn

cal components whose performance rates (levels) degrade with wear. Consequently, there are several intermediate states corresponding to the development of the wear and tear. Furthermore, as electrical parts, generators, converters and transformers still have intermediate states because of the backup. These features indicate that the traditional binary reliability model cannot perfectly characterize the DFIG.

A few works treated the DFIG as a multi-state system (MSS) and the reliability methods for MSS have been widely provided [18, 19, 22, 23]. Eryilmaz [8] presented a reliability method for a MSS with three-state components and applied it into wind energy. Xiao et al. have applied MSS model to many practical problems, and achievements have been made. [31, 32, 33]. UGF are considered as an convenient method for reliability assessment of MSS. Levitin has done a lot of research about UGF [16]. Nevertheless, the uncertainty exists in reliability assessment and cannot be neglected. Due to the limited data and/or vague judgments from experts, obtaining the accuracy value of the performance rates (levels) and probabilities of MSSs are difficult and inevitably contains epistemic uncertainty. Fuzzy set theory can well address the problems caused by the epistemic uncertainty. Huang et al. [11, 12, 13] developed a suit of reliability evaluation algorithms based on the fuzzy set theory. Wu [27] developed a fuzzy Bayesian method and proposed a new method to create the fuzzy Bayes point estimator of reliability. Ding and Lisnianski [6] combined the fuzzy set theory with the UGF technique; then, the fuzzy UGF (FUGF) method was proposed. Liu and Huang [21] further justified the FUGF method and introduced the Markov chain to the FUGF method. Lisnianski et al. [20] proposed an MSS reliability analysis and optimization method based on FUGF. Li et al. [17] provided an improved FUGF method for reliability assessment of MSS under aleatory and epistemic uncertainties. Gao et al. [10] performed dynamic fuzzy reliability analysis for MSS based on UGF. Dong et al. [7] extended the FUGF method for reliability assessment of uncertain MSS. Gao and Zhang [9] proposed a novel reliability analysis method for fuzzy MSS considering correlation based on UGF. The fuzzy theory and reliability analysis of MSS system have also developed recently [15, 24]. Jaiswal et al. [14] proposed Reliability analysis method for non-repairable weighted k-out-of-n system based on belief UGF. Cui et al. [5] presented a reliability model for aircraft actuation system based on power transfer efficiency. Qin et al. [26] proposed a combined method for reliability analysis of MSS of minor-repairable components. Negi and Singh [25] provided the fuzzy reliability evaluation method of linear m-consecutive weighted-k-out-of-r-from-n: F systems. Chen et al. [2] performed the reliability analysis and optimization of equal load-sharing k-out-of-n phased-mission systems.

In this paper, we consider a typical type of wind turbine generator systems, i.e., the DFIG, and model the DFIG as an MSS. As the DFIG is a system with high reliability and few test/event data, traditional reliability assessment methods based on large amount of failure data with rigorous statistical models are incapable of handling such a challenge. Moreover, specifying the component states of the DFIG, such as the states of the blades and gearboxes, often relies on experts' knowledge. Due to the vague judgements of experts, the determination of component states often contains epistemic uncertainty and it is suitable to be modelled as fuzzy numbers [21]. In this work, first, the fuzzy states of the DFIG systems are defined and the entire system state is calculated based on system structure function. Secondly, the performance rates and probabilities of all components' states are determined as triangular fuzzy number (TFN) based on experts' experiences. TFNs are chosen rather than other types of fuzzy numbers due to the easy concept and wide applications to reliability engineering. what's more, if the imprecise component state probability elicited by experts is naturally modelled by the TNFs, the experts only have to decide the most possible values of component state probability and the uncertainty associated with this decision. For other types of fuzzy numbers, such as Trapezoidal fuzzy number, the experts have to decide at least four values associated with the component state probability. It, therefore, introduces additional challenges to the expert elicitation process. Finally, the reliability assessment of the DFIG based on the FUGF is conducted.

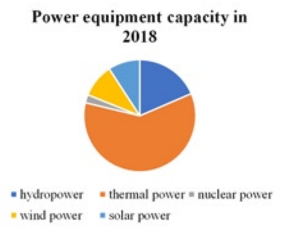
The remainder of this paper is organized as follows. Section 2 introduces a brief overview of the turbine generators. The introduction on the MSS and the UGF are given in Section 3. Section 4 conducts the fuzzy reliability assessment for the DFIG system based on FUGF. Finally, a brief conclusion is given in Section 5.

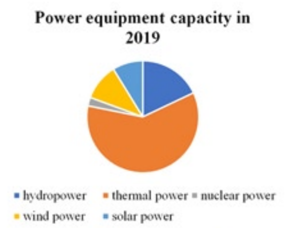
2. Overview of Wind Turbine Generators

2.1. Background and Structure of Wind Turbine Generators

As a significant new energy, wind power plays an indispensable role in both industry and our daily lives. In China, for example, wind power increases quickly in recent years, and it is ranked the third in the country's power equipment capacity. More information about Chinese power equipment capacity [3, 4] is shown in Fig. 1.

The wind turbine generator is the vital device to convert wind energy into electric energy. According to the output capacity of wind turbine generations, wind turbine generations can be divided into small, medium, large, and megawatt-scale. With the increase of the capacity, double fed wind induction generator (the DFIG) gradually becomes the mainstream of wind turbine generation market due to its good performance and operation stability. As for DFIG, there are five main parts, i.e., the blade, the gearbox, the generator, the converter and the transformer. The structure of DFIG is shown in Fig. 2. The blade can rotate with the wind and then the torque forming. This is the first step that the wind power transforms into mechanical energy. The forming torque will be transmitted to the gearbox for acceleration. The output shaft of gearbox with high-speed rotating is connected with the generator and then the mechanical energy will be transmitted into electric energy. The converter of DFIG is used to excite the rotor of the DFIG. The amplitude, frequency, and phase of the output voltage at the stator side of the DFIG are the same as those of the grid. Without the converter, the generator cannot work normally [30].





Type	Unit	Absolute amount	Increment
Hydropower	MWe	35226	2.5
Thermal power	MWe	114367	3.0
Nuclear power	MWe	4466	24.7
Wind power	MWe	18426	12.4
Solar power	MWe	17463	33.9
Total	MWe	189967	6.5

Type	Unit	Absolute amount	Increment
Hydropower	MWe	35640	1.1
Thermal power	MWe	119500	4.1
Nuclear power	MWe	4874	9.1
Wind power	MWe	21005	14
Solar power	MWe	20468	17.4
Total	MWe	201266	5.8

Fig. 1. Power equipment capacity of China in 2018 and 2019

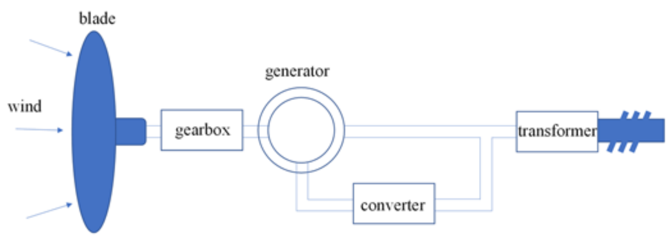


Fig. 2. The structure of DFIG

2.2. Reliability Modeling of a Wind Turbine Generator

According to the physical connections of the five components in DFIG, the reliability block diagram of a DFIG is shown in Fig. 3. In the DFIG, two separate systems with a generator and converter connected in series are used as redundancy. They are connected in series with the blade, the gearbox and the transformer.

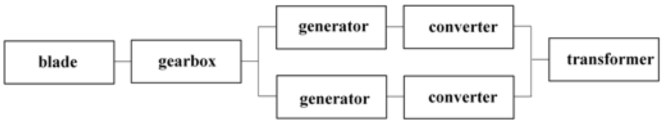


Fig. 3. Reliability block diagram of the DFIG.

The components of a DFIG can be categorized into two types, i.e., the mechanical type and the electrical type. The former includes the blade and the gearbox whose failures are mainly caused by wear. Whereas, the latter includes the generator, the converter and the transformer, whose failures are mainly caused by the damage of IGBT modules. For the former type of components, the mechanical performance will deteriorate into different levels with the development of wear and tear. When the performance reaches a certain threshold, components will be failure. For the latter type of components, the damage of IGBT modules can also make the degradation of performance due to the backup. Therefore, the system of the DFIG can be considered as an MSS whose components have multi-state as well. According to the experts' experiences, the states of every component of the DFIG is defined and given in Table 1. As we can see from Table 1, there are 4 states of blade and gearbox, 3 states of generator, convertor and transformer.

Table 1. States definition of DFIG.

Component	State division
Blade	perfect, mild wear, severe wear, failure (4 states)
Gearbox	perfect, mild wear, severe wear, failure (4 states)
Generator	perfect, middle, failure (3 states)
Converter	perfect, middle, failure (3 states)
Transformer	perfect, middle, failure (3 states)

Based on the structure function of the entire system, there are totally 482 states of the DFIG. Thus, it is difficult to accurately assess the system state parameters and an efficient method for reliability assessment is strongly needed.

3. UGF-Based Reliability Assessment of MSS

3.1. Overview of MSS

For an MSS, it could have a finite number of performance rates (levels). For each component, they could have a finite number of performance rates (levels) as well. In order to conduct the reliability assessment of an MSS, the characteristics of its components should be determined first. Components can have different states with corresponding performance rates (levels). The performance rates (levels) of every states of any components can be represented as:

$$\mathrm{g}_j = \left\{ g_{j1}, \cdots, g_{ji_j}, \cdots, g_{jk_j} \right\}, \tag{1}$$

where i_j indicates $i\mathrm{th}(1 \le i \le k_j)$ state of component and g_{i_j} is the performance rate (level) of j. Then the probabilities associated with different states of component *j* can be represented as:

$$\mathrm{p}_j = \left\{ p_{j1}, \cdots, p_{ji_j}, \cdots, p_{jk_j} \right\} \tag{2}$$

After determining the performance rates and corresponding probabilities, the probability distribution (PD) of the system can be determined if the system structure function $\phi(\bullet)$ is known. The probability of system state i can be calculated as follows:

$$p_i = \prod_{j=1}^{n} p_{ji_j} \tag{3}$$

The performance rate (level) of MSS for state i is:

$$g_i = \phi(g_{1i_1}, \cdots, g_{ni_n}) \tag{4}$$

The PD of the MSS can be represented as:

$$g_i = \phi(g_{1i_1}, \cdots, g_{ni_n}),\ p_i = \prod_{j=1}^{n} p_{ji_j} \tag{5}$$

3.2. UGF Method

The UGF is an effective method to conduct the reliability assessment of MSSs. Boolean model, stochastic process method, Monte Carlo simulation and UGF method are common method used for reliability analysis of MSS. In engineering practice, the UGF method can be applied to the system with complex structure and function, meanwhile, the calculation is small and the implementation is flexible. Most importantly, reliability assessment via the UGF method can be done by decomposing the calculation of system UGF into a combination of two component UGF. It, therefore, dramatically reduces the computational burden of system reliability assessment for complex systems with many components. As the performance rates and PD of the MSS have been determined, the $z-\mathrm{transform}$ of random variable $\mathrm{g}_j = \{g_{j1}, \cdots, g_{ji_j}, \cdots, g_{jk_j}\}$, $\mathrm{p}_j = \{p_{j1}, \cdots, p_{ji_j}, \cdots, p_{jk_j}\}$ is defined as:

$$u(z) = \sum_{i=1}^{k_j} p_{ji_j} \cdot z^{g_{ji_j}} \ . \tag{6}$$

Equation (6) represents the PD of the component j . This form is the UGF representation of multi-state component j . The output PD with $z-\mathrm{transform}$ representation of the entire system can be represented as:

$$\begin{aligned} U(\mathrm{z}) &= \Omega_\phi \left\{ u_1(z), \cdots, u_n(z) \right\} \\ &= \Omega_\phi \left\{ \sum_{i_1=1}^{k_1} p_{1i_1} \cdot z^{g_{1i_1}}, \cdots, \sum_{i_n=1}^{k_n} p_{ni_n} \cdot z^{g_{1i_1}} \right\}, \\ &= \sum_{i_1}^{k_1} \sum_{i_2}^{k_2} \cdots \sum_{i_n}^{k_n} \left(\prod_{j=1}^{n} p_{ji_j} \cdot z^{\phi(g_{1i_i}, \cdots, g_{ni_n})} \right) \end{aligned} \tag{7}$$

where Ω_ϕ is a general composition operator. The UGF method is based on the general composition operator and individual universal $z-\mathrm{transform}$ representations. Therefore, the PD of the MSS can be easily obtained through the PDs of each component if the structure function $\phi(\cdot)$ is known. The structure function $\phi(\cdot)$ is defined according to the structure of the system. A system with different structures, such as series, parallel, series-parallel or bridge structures, will have different $\phi(\cdot)$. The states of an MSS can be divided into two subsets depending on whether the state is acceptable by the system function. Whether a state is accepted or not depends on the system demand w . Suppose that the index $r_i : r_i = g_i - w$, the state *i* is an acceptable state if and only if $r_i \geq 0$. The availability of an MSS is the probability the system staying in the subset of acceptable states:

$$A(w) = \sum_{r_i \geq 0} p_i = \sum_{i=1}^{K} p_i \cdot \alpha_i \ , \tag{8}$$

where:

$$\alpha_i = \begin{cases} 1, \ r_i \geq 0 \\ 0, \ r_i < 0 \end{cases} . \tag{9}$$

Herein, a subsystem of the DFIG is taken as an example to illustrate the UGF-based reliability assessment for MSSs.

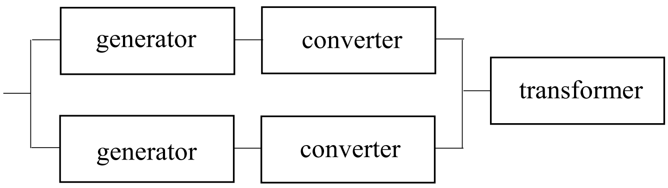


Fig. 4. Structure of the subsystem of DFIG

As shown in Fig. 4, there is a subsystem of DFIG with five components. This subsystem can be treated as a flow transmission system whose performance rate (level) is defined by their transmission capacity. Suppose that there are 3 states of generator (component 1) and converter (component 2) and 2 states of the transformer (component 3). The performance rates (levels) of the states of generator are $g_{11} = 1.7, g_{12} = 1.2, g_{13} = 0.5$ with the corresponding probabilities being $p_{11} = 0.7, p_{12} = 0.2, p_{13} = 0.1$, respectively. The performance rates (levels) of the states of the converter are $g_{21} = 0.8, g_{22} = 0.2, g_{23} = 0$ and the corresponding probabilities are $p_{21} = 0.4, p_{22} = 0.3, p_{23} = 0.3$. The performance rates (levels) of the states of transformer are $g_{31} = 1, g_{32} = 0$ and the corresponding probabilities $p_{31} = 0.5, p_{32} = 0.5$. The UGF for each component based on the PD is defined as:

$$u_1(z) = p_{11} \cdot z^{g_{11}} + p_{12} \cdot z^{g_{12}} + p_{13} \cdot z^{g_{13}} = 0.7 \cdot z^{1.7} + 0.2 \cdot z^{1.2} + 0.5 \cdot z^{0.5},$$
$$u_2(z) = p_{21} \cdot z^{g_{21}} + p_{22} \cdot z^{g_{22}} + p_{23} \cdot z^{g_{23}} = 0.4 \cdot z^{0.8} + 0.3 \cdot z^{0.2} + 0.3 \cdot z^{0},$$
$$u_3(z) = p_{31} \cdot z^{g_{31}} + p_{32} \cdot z^{g_{32}} = 0.5 \cdot z^{1} + 0.5 \cdot z^{0}.$$

According to the structure as shown in Figure 4, the system structure function is expressed as:

$$\begin{aligned} &\phi(G_1(t),\ G_2(t),\ G_3(t)) \\ &= \phi_s\left\{\phi_p\left[\phi_s(G_1(t),\ G_2(t)),\phi_s(G_1(t),\ G_2(t))\right],G_3(t)\right\} \\ &= \min\left\{\left[\min(G_1(t),\ G_2(t)) + \min(G_1(t),\ G_2(t))\right],G_3(t)\right\} \end{aligned} ,$$

and the PD of the entire system can be obtained as:

$$\begin{aligned} &\Omega_\phi(u_1(z),\ u_2(z),\ u_3(z)) \\ &= \Omega_{\phi s}\left\{\Omega_{\phi p}\left[\Omega_{\phi s}(u_1(t),\ u_2(t)),\Omega_{\phi s}(u_1(t),\ u_2(t))\right],u_3(t)\right\} \\ &= \Omega_{\phi s}\begin{Bmatrix} 0.1296 \cdot z^{1.6} + 0.0288 \cdot z^{1.3} + 0.2176 \cdot z^{1} + 0.216 \cdot z^{0.8} + 0.024 \cdot z^{0.7} \\ +0.024 \cdot z^{0.5} + 0.09 \cdot z^{0.4} + 0.18 \cdot z^{0.2} + 0.09 \cdot z^{0}, 0.5 \cdot z^{1} + 0.5 \cdot z^{0} \end{Bmatrix} \\ &= 0.188 \cdot z^{1} + 0.108 \cdot z^{0.8} + 0.012 \cdot z^{0.7} + 0.012 \cdot z^{0.5} + 0.045 \cdot z^{0.4} + 0.09 \cdot z^{0.2} + 0.545 \cdot z^{0} \end{aligned} .$$

Therefore, the system availability is calculated as:

$$\begin{aligned} A(0.5) &= \sum_{i=1}^{4} \mathrm{p_i} \cdot \alpha_i = 0.188 + 0.108 + 0.012 + 0.01 \\ &= 0.32 \end{aligned} .$$

The result of 0.32 can be considered as the system availability corresponding to the demand w=0.5, and the reliability of the system can be assessed if the reliability demand is a set.

4. FUGF Method for Reliability Assessment of the DFIG

4.1. FUGF

Fuzzy reliability theory is a combination of fuzzy mathematics and reliability theory. Conventional UGF technique is based on two fundamental assumptions. Firstly, the probabilities of each state of each component can be fully characterized by probability measures. Secondly, the performance rate of each component can be precisely determined. However, since the performance rates and probabilities cannot be obtained precisely in practical engineering, the FUGF technique is developed. Therefore, the values in UGF cannot be represented as crisp numbers and the values can be considered around a crisp number. In this situation, the fuzzy set and fuzzy number are proposed to describe such epistemic uncertainty.

A fuzzy number is different from a crisp number because it is a subset defined by its membership function. For example, $\tilde{X}$ is a fuzzy subset, and is defined by its membership function $\mu_{\overline{X}}(x): U \to [0,1]$. The values of $\mu_{\overline{X}}(x)$ are in the range of 0 to 1, and the value of $\mu_{\overline{X}}(x)$ indicates the probability that the fuzzy number can be obtained as a specific value x. There are different kinds of fuzzy numbers with different kind of membership functions. In this paper, the TFN is considered. The membership function of a typical TFN parameterized by the triplet is defined as:

$$\mu_{\overline{X}}(x) = \begin{cases} \dfrac{x-a}{b-a}, \ a \leq x \leq b \\ 1, \qquad x = b \\ \dfrac{x-c}{b-c}, \ b < x \leq c \\ 0, \qquad \text{otherwise} \end{cases} \tag{10}$$

And the function can be plotted as Fig. 5.

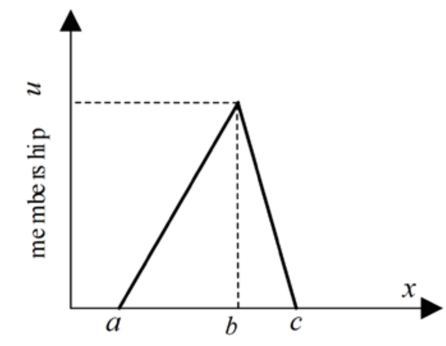


Fig. 5. Membership function of TFN

If fuzzy values exist in the UGF, it can be considered as FUGF. In this paper, both the performance rates and the probabilities are treated as fuzzy numbers. Furthermore, all the fuzzy numbers in this paper are considered as TFNs.

For a fuzzy MSS with n components, the component $j (1 \leq j \leq n)$ can have k_j different states, the corresponding PD can be represented as ordered fuzzy sets $\tilde{\mathrm{g}}_{\mathrm{j}} = \left\{\tilde{g}_{j1}, \cdots, \tilde{g}_{ji_j}, \cdots, \tilde{g}_{jk_j}\right\}$ and

$\tilde{p}_j = \left\{ \tilde{p}_{j1}, \cdots, \tilde{p}_{ji_j}, \cdots, \tilde{p}_{jk_j} \right\}$, so the fuzzy performance rates (levels) and probabilities of each state are:

$$\begin{cases} \tilde{p}_{ji_j} = \left\{ p_{ji_j}, \mu_{\tilde{p}_{ji_j}}(p_{ji_j}) \mid p_{ji_j} \in P_{ji_j} \right\} \\ \tilde{g}_{ji_j} = \left\{ \tilde{g}_{ji_j}, \mu_{\tilde{g}_{ji_j}}(g_{ji_j}) \mid g_{ji_j} \in G_{ji_j} \right\} \end{cases} \tag{11}$$

where $\mu_{\tilde{p}_{ji_j}}$ and $\mu_{\tilde{g}_{ji_j}}$ are membership function of $\tilde{p}_{ji_j}$ and $\tilde{g}_{ji_j}$, P_{ji_j} and G_{ji_j} are collection of objects denoted by $\tilde{p}_{ji_j}$ and $\tilde{g}_{ji_j}$, respectively.

The operation of fuzzy number follows the extension principle, the performance of system state i can be evaluated as:

$$\begin{aligned} \tilde{g}_i &= \tilde{\phi}(\tilde{g}_{1i_1}, \cdots \tilde{g}_{ji_j}, \cdots, \tilde{g}_{ni_n}) \\ &= \left\{ g_i, \mu_{\tilde{g}_i}(g_i) \mid g_i = \phi(g_{1i_1}, \cdots, g_{ji_j}, \cdots, g_{ni_n}), g_{ji_j} \in G_{ji_j} \right\} \end{aligned} \tag{12}$$

where $\mu_{\tilde{g}_i}(g_i) = \sup_{\phi(g_{1i_1}, \cdots, g_{ji_j}, \cdots, g_{ni_n})} \min \left\{ \mu_{g_{1i_1}}, \cdots, \mu_{g_{ji_j}} \right\}$, and $\phi(g_{1i_1}, \cdots, g_{ji_j}, \cdots, g_{ni_n})$ is the structure function of FMSS.

The probability of system state i represented by fuzzy numbers can be calculated as:

$$\tilde{p}_i = \left\{ p_i, \mu_{\tilde{p}_i}(p_i) \mid p_i = \prod_{j=1}^{n} p_{ji_j}, p_{ji_j} \in P_{ji_j} \right\} \tag{13}$$

where $\mu_{\tilde{p}_i}(p_i) = \sup_{p_i = \prod_{j=1}^{n} p_{ji_j}} \min \left\{ \mu_{p_{1i_1}}, \cdots, \mu_{p_{ji_j}} \right\}$.

The PD of a FMSS can be calculated as:

$$\begin{aligned} U(z) &= \Omega_{\phi}\left(\sum_{i_1=1}^{k_1} \tilde{p}_{1i_1} \cdot z^{\tilde{g}_{1i_1}}, \cdots, \sum_{i_n=1}^{k_n} \tilde{p}_{ni_n} \cdot z^{\tilde{g}_{1i_1}} \right) \\ &= \sum_{i_1}^{k_1} \sum_{i_2}^{k_2} \cdots \sum_{i_n}^{k_n} \left(\tilde{p}_i \cdot z^{\phi(\tilde{g}_{1i_1}, \cdots, \tilde{g}_{ni_n})} \right) \\ &= \sum_{i_1}^{k_1} \sum_{i_2}^{k_2} \cdots \sum_{i_n}^{k_n} \left(\tilde{p}_i \cdot z^{\tilde{g}_i} \right). \end{aligned} \tag{14}$$

Since system demand is represented as a fuzzy number, the availability assessment for a fuzzy MSS is re-defined in this paper. If the performance rate (level) g for the state i is represented as a TFN parameterized by a triplet (a,b,c) and the system demand w is represented as a TFN parametrized by a triplet (x,y,z), there would be different kinds of relationship between them.

If $a \geq x$, state i is a reliable state.

If $x \geq c$, state i is a failure state.

If there is an overlapping between (a,b,c) and (x,y,z), $|ar_i|_{\text{rel}}$ is defined to obtain the availability. The availability of a FMSS can be represented as:

$$\tilde{A}(\tilde{w}) = \sum_{I=I}^{k} \tilde{p}_i \cdot |ar_i|_{\text{rel}}, \tag{15}$$

where $|ar_i|_{\text{rel}}$ is the relative cardinality of fuzzy set $\tilde{ar_i}$ and $\tilde{ar_i} = \left\{ ar_i, \mu(ar_i) \mid \mu(ar_i) = \mu(r_i), ar_i \in AR_i \right\}$. $|ar_i|_{\text{rel}}$ can be obtained by the following equations:

$$\begin{cases} |r_i| = \sum_{r_i \in R_i} \mu_{\tilde{r}_i}(r_i) \\ |ar_i| = \sum_{ar_i \in AR_i} \mu_{\tilde{ar_i}}(ar_i) \\ |ar_i|_{\text{rel}} = |ar_i| / |r_i| \end{cases} \tag{16}$$

where $|ar_i|_{\text{rel}}$ is the relative cardinality of fuzzy set $\tilde{ar_i}$, and $AR_i = \left\{ r_i \in R \mid r_i \geq 0 \right\}$, $\tilde{ar_i} = \left\{ ar_i, \mu(ar_i) \mid \mu(ar_i) = \mu(r_i), ar_i \in AR_i \right\}$.

4.2. Reliability Assessment of the DFIG by FUGF

From the state definitions of DFIG in Table 1, there are 4 states of the blade and the gearbox, and 3 states of the generator, the converter and the transformer, respectively. The degradation forms of components are different, for instance, the blade will have a slower speed of rotation but the gearbox will have a lower speed of the output shaft during degradation. Due to limited reliability testing resources (e.g., time, budget, manpower), the amount of collected reliability-related data from the components of the DFIG are extremely small. It, therefore, becomes difficult to estimate the precise values of the state probabilities of the DFIG and its components [28], [29]. Alternatively, imprecise information with respect to the DFIG and its components states, i.e., the performance rates (levels), and the corresponding state probabilities can be gathered from experts. In this work, the performance rates (levels) of all components are treated as TFNs under the fuzzy set theory, as tabulated in Table 2. The data in Table 2 are collected from real industry according to cooperation with wind turbine enterprises.

Based on the given values, the reliability assessment based on FUGF can be conducted as follows:

$$\begin{aligned} u_1(z) &= \tilde{p}_{11} \cdot z^{\tilde{g}_{11}} + \tilde{p}_{12} \cdot z^{\tilde{g}_{12}} + \tilde{p}_{13} \cdot z^{\tilde{g}_{13}} + \tilde{p}_{14} \cdot z^{\tilde{g}_{14}} = (0.72, 0.76, 0.77) \cdot z^1 \\ &+ (0.1, 0.11, 0.13) \cdot z^{(0.7, 0.8, 0.85)} + (0.06, 0.08, 0.1) \cdot z^{(0.45, 0.5, 0.6)} + (0.03, 0.05, 0.06) \cdot z^0, \\ u_2(z) &= \tilde{p}_{21} \cdot z^{\tilde{g}_{21}} + \tilde{p}_{22} \cdot z^{\tilde{g}_{22}} + \tilde{p}_{23} \cdot z^{\tilde{g}_{23}} + \tilde{p}_{24} \cdot z^{\tilde{g}_{24}} = (0.71, 0.72, 0.75) \cdot z^1 \\ &+ (0.2, 0.21, 0.23) \cdot z^{(0.75, 0.78, 0.8)} + (0.03, 0.05, 0.08) \cdot z^{(0.35, 0.4, 0.48)} + (0.01, 0.02, 0.03) \cdot z^0, \\ u_3(z) &= \tilde{p}_{31} \cdot z^{\tilde{g}_{31}} + \tilde{p}_{32} \cdot z^{\tilde{g}_{32}} + \tilde{p}_{33} \cdot z^{\tilde{g}_{33}} = (0.77, 0.83, 0.86) \cdot z^1 \\ &+ (0.12, 0.15, 0.2) \cdot z^{0.7} + (0.02, 0.06, 0.09) \cdot z^0, \\ u_4(z) &= \tilde{p}_{41} \cdot z^{\tilde{g}_{41}} + \tilde{p}_{42} \cdot z^{\tilde{g}_{42}} + \tilde{p}_{43} \cdot z^{\tilde{g}_{43}} = (0.81, 0.87, 0.92) \cdot z^1 \\ &+ (0.04, 0.1, 0.12) \cdot z^{0.6} + (0.01, 0.03, 0.06) \cdot z^0, \\ u_5(z) &= \tilde{p}_{51} \cdot z^{\tilde{g}_{51}} + \tilde{p}_{52} \cdot z^{\tilde{g}_{52}} + \tilde{p}_{53} \cdot z^{\tilde{g}_{53}} = (0.76, 0.82, 0.84) \cdot z^1 \\ &+ (0.11, 0.13, 0.17) \cdot z^{0.55} + (0.03, 0.05, 0.08) \cdot z^0. \end{aligned}$$

According to Fig. 3, components 3 and 4 are connected in parallel. Therefore, the operator $\tilde{\Omega}_{\tilde{\phi}_P}$ is applied between $u_1(z)$ and $u_2(z)$:

$$\begin{aligned} &\tilde{\Omega}_{\tilde{\phi}_s}(u_3(z), u_4(z)) \\ &= (0.6237,\ 0.7221,\ 0.7912) \cdot z^1 + (0.0972,\ 0.1305,\ 0.1840) \cdot z^{0.7} \\ &+ (0.0356,\ 0.0980,\ 0.1272) \cdot z^{0.6} + (0.0261,\ 0.0878,\ 0.1626) \cdot z^0 \\ &\tilde{\Omega}_{\tilde{\phi}_P}[\tilde{\Omega}_{\tilde{\phi}_s}(u_3(z), u_4(z)), \tilde{\Omega}_{\tilde{\phi}_s}(u_3(z), u_4(z))] \\ &= (0.3890,\ 0.5214,\ 0.6260) \cdot z^2 + (0.1212,\ 0.1884,\ 0.2912) \cdot z^{1.7} \\ &= (0.0444,\ 0.1416,\ 0.2012) \cdot z^{1.6} + (0.0094,\ 0.0170,\ 0.0339) \cdot z^{1.4} \\ &= (0.0070,\ 0.0256,\ 0.0468) \cdot z^{1.3} + (0.0013,\ 0.0096,\ 0.0162) \cdot z^{1.2} \\ &= (0.0326,\ 0.1268,\ 0.2572) \cdot z^1 + (0.0050,\ 0.0230,\ 0.0598) \cdot z^{0.7} \\ &= (0.0018,\ 0.0172,\ 0.0414) \cdot z^{0.6} + (0.0007,\ 0.0077,\ 0.0264) \cdot z^0 \end{aligned}$$

Table 2. Performance rates and probabilities of each component of DFIG

Component (No.)	State (No.)	Performance rate	Probability
Blade (1)	Perfect (11)	1	(0.72, 0.76, 0.77)
	mild wear (12)	(0.7, 0.8, 0.85)	(0.1, 0.11, 0.13)
	severe wear (13)	(0.45, 0.5, 0.6)	(0.06, 0.08, 0.1)
	Failure (14)	0	(0.03, 0.05, 0.06)
Gearbox (2)	Perfect (21)	1	(0.71, 0.72, 0.75)
	mild wear (22)	(0.75, 0.78, 0.80)	(0.20, 0.21, 0.23)
	severe wear (23)	(0.35, 0.4, 0.48)	(0.03, 0.05, 0.08)
	Failure (24)	0	(0.01, 0.02, 0.03)
Generator (3)	Perfect (31)	1	(0.77, 0.83, 0.86)
	Middle (32)	0.7	(0.12, 0.15, 0.2)
	Failure (33)	0	(0.02, 0.06, 0.09)
Converter (4)	Perfect (41)	1	(0.81, 0.87, 0.92)
	Middle (42)	0.6	(0.04, 0.1, 0.12)
	Failure (43)	0	(0.01, 0.03, 0.06)
Transformer (5)	Perfect (51)	1	(0.76, 0.82, 0.84)
	Middle (52)	0.55	(0.11, 0.13, 0.17)
	Failure (53)	0	(0.03, 0.05, 0.08)

The FUGF of the entire system can be obtained as follows:

$$\tilde{\Omega}_S = \tilde{\Omega}_{\tilde{\phi}_S}\left\{u_1(z),\ u_2(z),\ \tilde{\Omega}_{\tilde{\phi}_P}[\tilde{\Omega}_{\tilde{\phi}_S}(u_3(z),\ u_4(z)),\tilde{\Omega}_{\tilde{\phi}_S}(u_3(z),\ u_4(z))],\ u_5(z)\right\} = \sum_{i=1}^{432}\tilde{p}_i \cdot z^{\tilde{g}_i}$$

Since there are 432 states of the DFIG system, it is difficult to list all the states. Thus, the successful states are considered to conduct the availability assessment. Let $\tilde{\Omega}_S$ be the FUGF of the acceptable states with the system demand is $(0.78,\ 0.85,\ 0.92)$, then we have:

$$\tilde{\Omega}_S = (0.2350,\ 0.4623,\ 0.7143)\cdot z^1 + (0.0326,\ 0.0669,\ 0.1206)\cdot z^{(0.8,\ 0.83,\ 0.85)}$$

$$|\, ar_2\,|_{\text{rel}} = 1/7 = 0.1429$$

$$\tilde{A}(\tilde{w}) = (0.2350,\ 0.4623,\ 0.7143) + (0.0326,\ 0.0669,\ 0.1206) * 0.1429 = (0.2397,\ 0.4719,\ 0.7315)$$

5. Conclusions

In this paper, the reliability assessment of the DFIG, a typical wind turbine generator, is conducted under the fuzzy set theory. The DFIG, which consists of a blade, a gearbox, a generator, a converter, and a transformer is treated as an MSS. The FUGF method is used to evaluate the reliability of the DFIG with fuzzy states and probabilities. Firstly, the reliability block diagram of DFIG is built according to the system structure function. Secondly, the component states are defined. Specifically, there are four (perfect, mild wear, severe wear and failure) states for the blade, 4 states (perfect, mild wear, severe wear and failure) for the gearbox, 3 states (perfect, middle and failure) for the generator, 3 states (perfect, middle and failure) for the converter, and 3 states (perfect, middle and failure) for transformer. Finally, the FUGF method is used to calculate the fuzzy availability of the entire DFIG system based on the reliability block diagram, fuzzy states and the corresponding probabilities. The results show that given the system demand $(0.78,\ 0.85,\ 0.92)$, the availability of the DFIG system is $(0.2397,\ 0.4719,\ 0.7315)$. If the system demand increases to a higher level, the availability of system will decrease and vice versa. Triangular fuzzy number is a special category of trapezoidal fuzzy number, which is the most widely studied fuzzy number. Most fuzzy concepts and fuzzy information in real life, especially some fuzzy judgments of decision-makers or experts' experience, can be expressed by triangular fuzzy numbers. It is noteworthy that the proposed constrained optimization model for reliability assessment is not restricted to the TNF. It can be readily implemented to other types of fuzzy numbers because at any cut levels, we can find the interval of the component state probability of any type of fuzzy numbers. Therefore, the proposed method is a generalized method for reliability assessment under fuzzy set theory. However, for reliability assessment of MSS, UGF method is convenient, but it is not equal to that UGF method is the most accurate one. In the future, we need to further compare the accuracy of the results with other methods. This is the direction we will focus on in the future.

Acknowledgement

The authors extend sincere gratitude to the National Natural Science Foundation of China for financial support under contract number 51775090.

References

1. Carroll J, McDonald A, Mcmillan D. Reliability comparison of wind turbines with DFIG and PMG drive trains. In 2015 IEEE Power & Energy Society General Meeting, Denver, CO: IEEE 2015; 663-670, https://doi.org/10.1109/PESGM.2015.7286449.
2. Chen JK, He Y, Wei W. Reliability analysis and optimization of equal load-sharing K-out-of-N phased-mission systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(2): 250-259, https://doi.org/10.17531/ein.2015.2.12.
3. China National Energy Administration. National electric power industry statistics in 2019. 2020.
4. China National Energy Administration. National electric power industry statistics in 2018. 2019.
5. Cui XY, Li TY, Wang S P, Shi J, Ma Z H. Reliability modeling based on power transfer efficiency and its application to aircraft actuation system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(2): 282-296, https://doi.org/10.17531/ein.2020.2.11.

6. Ding Y, Lisnianski A. Fuzzy universal generating functions for multi-state system reliability assessment. Fuzzy Sets and System 2007; 159: 307-324, https://doi.org/10.1016/j.fss.2007.06.004.
7. Dong W, Liu S, Zhang Q, Mierzwiak R, Fang Z, Cao Y. Reliability assessment for uncertain multi-state systems: an extension of fuzzy universal generating function. International Journal of Fuzzy Systems 2019; 21, 945-953, https://doi.org/10.1007/s40815-018-0552-x.
8. Eryilmaz S. Reliability analysis of multi-state system with three-state components and its application to wind energy. Reliability Engineering & System Safety 2017; 172: 58-63, https://doi.org/10.1016/j.ress.2017.12.008.
9. Gao H, Zhang X. A novel reliability analysis method for fuzzy multi-state systems considering correlation. IEEE Access 2019; 7: 153194-153204, https://doi.org/10.1109/ACCESS.2019.2948497.
10. Gao P, Xie L, Hu W, Liu C, Feng J. Dynamic fuzzy reliability analysis of multistate systems based on universal generating function. Mathematical Problems in Engineering 2018; 2018: 1-8, https://doi.org/10.1155/2018/6524629.
11. Huang HZ. Reliability analysis method in the presence of fuzziness attached to operation time. Microelectronics Reliability 1995; 35: 1483-1487, https://doi.org/10.1016/0026-2714(94)00173-L.
12. Huang HZ, Tong X, Zuo MJ. Posbist fault tree analysis of coherent systems. Reliability Engineering & System Safety 2004; 84: 141-148, https://doi.org/10.1016/j.ress.2003.11.002.
13. Huang HZ, Zuo MJ, Sun Z. Bayesian reliability analysis for fuzzy lifetime data. Fuzzy Sets and System 2006; 157: 1674-86, https://doi.org/10.1016/j.fss.2005.11.009.
14. Jaiswal N, Negi S, Singh S. Reliability analysis of non-repairable weighted k-out-of-n system using belief universal generating function. International Journal of Industrial and Systems Engineering 2018; 28: 300-318, https://doi.org/10.1504/IJISE.2018.089741.
15. Khaniyev T, Baskir M B, Gokpinar F, Mirzayev F. Statistical distribution and reliability functions with type-2 fuzzy parameters. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 268-274, https://doi.org/10.17531/ein.2019.2.11.
16. Levitin G. The universal generating function in reliability analysis and optimization. London: Springer, 2005.
17. Li Y, Ding Y, Zio E. Random fuzzy extension of the universal generating function approach for the reliability assessment of multi-state systems under aleatory and epistemic uncertainties. IEEE Transactions on Reliability 2014; 63: 13-25, https://doi.org/10.1109/TR.2014.2299031.
18. Li YF, Huang HZ, Mi J, Peng W, Han X. Reliability analysis of multi-state systems with common cause failures based on Bayesian network and fuzzy probability. Annals of Operations Research 2019; https://doi.org/10.1007/s10479-019-03247-6.
19. Li YF, Liu Y, Huang T, Huang HZ, Mi J. Reliability assessment for systems suffering common cause failure based on Bayesian networks and proportional hazards model. Quality and Reliability Engineering International 2020; 36(7): 2509-2520, https://doi.org/10.1002/qre.2713.
20. Lisnianski A, Frenkel I, Ding Y. Multi-State System Reliability Analysis and Optimization for Engineers and Industrial Managers. Spring Verlag: Berlin, New York, 2010, https://doi.org/10.1007/978-1-84996-320-6.
21. Liu Y, Huang HZ. Reliability assessment for fuzzy multi-state systems. International Journal of Systems Science 2009; 41: 365-379, https://doi.org/10.1080/00207720903042939.
22. Mi J, Li YF, Liu Y, Yang YJ, Huang HZ. Belief universal generating function analysis of multi-state systems under epistemic uncertainty and common cause failures. IEEE Transactions on Reliability 2015; 64(4): 1300-1309, https://doi.org/10.1109/TR.2015.2419620.
23. Mi J, Li YF, Peng W, Huang HZ. Reliability analysis of complex multi-state system with common cause failure based on evidential networks. Reliability Engineering & System Safety 2018; 174: 71-81, https://doi.org/10.1016/j.ress.2018.02.021.
24. Mlynarski S, Pilch R, Smolnik M, Szybka J, Wiazania G. A method for rapid evaluation of k-out-of-n systems reliability. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 170-176, https://doi.org/10.17531/ein.2019.1.20.
25. Negi S, Singh S. Fuzzy reliability evaluation of linear m-consecutive weighted-k-out-of-r-from-n: F systems. International Journal of Computing Science and Mathematics 2019; 10: 606-621, https://doi.org/10.1504/IJCSM.2019.104027.
26. Qin JL, Niu YG, Li Z. A combined method for reliability analysis of multiple-state system of minor-repairable components. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(1): 80-88, https://doi.org/10.17531/ein.2016.1.11.
27. Wu H. Fuzzy reliability estimation using Bayesian approach. Computers and Industrial Engineering 2004; 46: 467-93, https://doi.org/10.1016/j.cie.2004.01.009.
28. Xiahou T, Liu Y. Reliability bounds for multi-state systems by fusing multiple sources of imprecise information. IISE Transactions 2020; 52(9): 1014-1031, https://doi.org/10.1080/24725854.2019.1680910.
29. Xiahou T, Zeng Z, Liu Y. Remaining useful life prediction by fusing experts' knowledge and condition monitoring information. IEEE Transactions on Industrial Informatics 2020, https://doi.org/10.1109/TII.2020.2998102.
30. Xiao H, Cao M. Joint Optimization of redundancy and preventive maintenance of a power grid system considering economic cost. Energy 2020; 201: 118470, https://doi.org/10.1016/j.energy.2020.118470.
31. Xiao H, Cao MH, Kou G, Yuan XJ. Optimal element allocation and sequencing of multi-state series systems with two levels of performance sharing. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2019, https://doi.org/10.1177/1748006X20947846.
32. Xiao H, Yi KX, Kou G, Xing LD. Reliability of a two-dimensional demand-based networked system with multistate components. Naval Research Logistics (NRL), 2020, 67(6): 453-468, https://doi.org/10.1002/nav.21922.
33. Xiao H, Zhang YY, Xiang YS, Peng R. Optimal design of a linear sliding window system with consideration of performance sharing. Reliability Engineering & System Safety, 2020, 198: 106900, https://doi.org/10.1016/j.ress.2020.106900.
34. Zhou D, Blaabjerg F, Lau M. Cost on reliability and production loss for power converters in the doubly fed induction generator to support modern grid codes. Electric Power Components and Systems 2016; 44: 152-164, https://doi.org/10.1080/15325008.2015.1102990.

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Zvolenský P, Barta D, Grenčík J, Droździel P, Kašiar Ľ. Improved method of processing the output parameters of the diesel locomotive engine for more efficient maintenance. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 315–323, http://doi.org/10.17531/ein.2021.2.11.

Improved method of processing the output parameters of the diesel locomotive engine for more efficient maintenance

Indexed by: Web of Science Group

Peter Zvolenský[a], Dalibor Barta[a], Juraj Grenčík[a], Pawel Droździel[b], Ľubomír Kašiar[c]

[a]University of Žilina, Faculty of Mechanical Engineering, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic
[b]Lublin University of Technology, Mechanical Engineering Faculty, ul. Nadbystrzycka 36, 20-618 Lublin, Poland
[c]Schaeffler Kysuce spol s.r.o., Dr. G. Schaefflera 1, 024 01 Kysucké Nové Mesto, Slovak Republic

Highlights

- Expert-based determination of the fault limit values of the engine parameters.
- Methodology for simpler display of fault parameters from the diagnostics.
- Improved diagnostics by faster, easier and more accurate fault evaluation.
- Increased maintenance effectiveness by shortening locomotive's downtime.

Abstract

Modernization of aged rolling stock is one of the possibilities to adapt it to the current requirements for better environmental friendliness and economy of railway transport. However, some vehicle upgrades lead to new failures that were not observed in the original vehicles. The cause is the so-called "hybrid design", built on a combination of original and selected new components. The aim of the work was to improve the situation with frequent failures and unavailability that occur on the modernized locomotive where a new diesel engine and new electronic control system was installed. Within the work, a simplified methodology for evaluating the outputs of diagnostic equipment was developped based on and applied to specific locomotive type and its diesel engine. The methodology resulted in a significant reduction of the time for assessing the condition of the vehicle's diesel engine and more effective maintenance. The paper also presents other possibilities in the analysis of big data in the maintenance of rolling stock e.g. using fuzzy logic.

Keywords

maintenance, locomotive, diesel engine, control parameters.

1. Introduction

The basic function/purpose of maintenance of railway vehicles is to keep them running safely and economically for expected life cycle. Therefore, fundamental is to do everything to avoid failures that could end with serious consequences, primarily derailment of a vehicle and train. European Union is aware of importance of good technical condition of rail vehicles and the role of maintenance. So it adapted extensive requirements and legislation in the area of rail safety. Conscience overview on legal requirements regarding railway transport safety in the European Union the practical solutions developed for railway operators as a part of the implementation of maintenance management systems can be found in [42]. Methodology for building a strategy of maintenance focused on safety of railway vehicles using RAMS (Reliability, Availability, Maintainability, Safety) analysis is described in [45] where on the example of a diesel locomotive it was found that this analysis enables proper classification of hazards, quantification of the frequency of occurrence of hazards and the adoption of the appropriate criteria for risk assessment of the created strategy. As stated also in [39], vehicle reliability is strongly linked to rail safety.

Maintenance plays an essential role in a system's life cycle. At the system level, the maintenance influences the reliability and availability of the system [7]. Achieving quality maintenance of any technical system, including railway technology, requires choosing the right maintenance strategy. However, we must realize that more comprehensive maintenance means higher life cycle costs, although this may not lead to a significant improvement in reliability. Measuring and assessing maintenance performance is critical to the competitiveness and future survival of any company providing production or services. As stated Mlynarski et al. the economic indicators of the operation process are one of the most important indicators of the use of vehicles in transport systems. This is because it is the operation management that largely determines the proper functioning of the entire business company [28].

2. Problematic of rail vehicle maintenance and modernization

Macián considers maintenance to be one of the largest expenditures for the transport companies together with fuel (or energy) costs and drivers (personnel) [24], which is, however, the most important one

E-mail addresses: P. Zvolenský - peter.zvolensky@fstroj.uniza.sk, D. Barta - dalibor.barta@fstroj.uniza.sk, J. Grenčík - juraj.grencik@fstroj.uniza.sk, P. Droździel - p.drozdziel@pollub.pl, Ľ. Kašiar - kasialbo@schaeffler.com

from the view of controllability, attending that fuel and labour costs are more externally driven (crude prices volatility, taxes, personnel policies and salaries, etc.) [23]. As in a free market, the optimal maintenance strategy can not only guarantee the availability of railway system but also have the best economic benefits [43].

The maintenance dealing with safety critical components is in particular concerning the wheelsets and parts connected with running gear in general, such as bearings, suspension etc. In this area the durability and prevention of any failure is the most important task. There are numerous standard and verified methods used in technical condition assessment and remaining utilization life prediction methods. With new diagnostics technology and computer support, extensive research is carried out.

In [21] a general reliability study using both classical and Bayesian semi-parametric degradation approaches for reliability analysis are presented. It is illustrated how degradation data can be modelled and analyzed to determine reliability to support preventive maintenance strategy based on a data-driven framework. With the proposed classical approach, both accelerated life tests and design of experiments technology are used to determine how each critical factor affects the prediction of performance, in this study demonstrated on a locomotive wheel-set reliability, being a safety critical component. Other safety critical components requiring higher attention and monitoring are axle bearings. Authors in [35, 51] present the prognostics and health monitoring concept in rail vehicles, specifically focused in bearing health state and remaining useful life. This concept is rapidly growing field of research with the aim of improving the reliability and availability of railway systems switching from time-based to event-driven maintenance policies.

Assessment of the reliability based on statistical methods is so far mostly used. Practical example can be found in the study [16] where statistical methods of quality management were used to identify the most problematic types of diesel locomotive equipment and specific causes of failures. Similar approach to reliability of power equipment on electric locomotives is described in [17], where failure occurrence and reliability analysis was done considering the negative impact of climatic factors on operation and technical condition of the rolling stock.

Increasing demands on the economy and safety of operation of various devices lead to the prediction of remaining service life (RUL). Most researchers devoted to improve the accuracy of the prediction results, and have investigated many effective methods for RUL prediction, including various neural networks (NN), support vector regression (SVR), stochastic process, and other methods. Ramezani et al. give a comprehensive summary to various methods [37]. The researchers form NASA Ames Research Center suggests that the perfectly and precisely prediction of engineering systems behavior is not possible in practical engineering applications due to prognostics uncertainty, and divided the sources of uncertainty into four categories: present state uncertainty, future uncertainty, modeling uncertainty and prediction method uncertainty [40].

In the literature, apart from the classical maintenance models [31], there are numerous maintenance models available on the required reliability level of an entire system [53], some of them based on the application of simulation methods [10, 29]. Some models include the possibility of partial maintenance [28], some others make use of additional inspection of object's technical state when it can be performed while the system is in actual operation [6]. A detailed and comprehensive classification of existing preventive renewal models is provided in [49]. In practical engineering, besides the randomness that can be modeled by probabilistic theory with probability distribution functions, epistemic uncertainty is another issue, caused by factors such as loss of information, limited knowledge, and inevitable man-made mistakes [15], which cannot be well explained by randomness and probabilistic models.

For the system maintenance and availability analysis, there are mathematical formulating and model-based analysis approaches. Garmabaki et al. presented the Multi-Attribute Utility Theory (MAUT), which used multiple objective functions to evaluate the cost and reliability of the maintenance optimization [11]. A gamma deterioration process was proposed by Meier-Hirmer et al., and it was applied to analyze the track maintenance [25]. Furthermore, the Maintenance Engineering Department of French National Railway Company (SNCF) introduced a formal method to estimate the maintenance strategy [3, 43].

Availability studies for degrading systems have been carried out by numerous researchers, but these are mainly based on Markov model using constant failure and repair rates, which is unrealistic in actual operating conditions. Markov model is a stochastic model which is used to model randomly changing systems over time. The basic assumption of a Markov Process is that the behavior of a system in each state is memory less which illustrates that the future evolution of the process depends only on the present state and not on the past sequence of traversed states prior to current state [19].

An interesting observation about relation between operation and maintenance is described in [46], where it is emphasized that maintenance strategy contributes higher efficiency of railway vehicles. One of new solutions that improve economy and thus effectiveness of maintenance is use of mobile maintenance points, as reported in [47].

The purpose of all the technical solutions in condition monitoring and reliability analysis after all is to create more effective maintenance system - maintenance planning and execution. Knowledge of distributions of times to failure is fundamental for maintenance planning [41]. The proposed methods for improved maintenance schedules and new algorithms for overhauls planning are defined in [20, 34].

From the investigation of the state-of-the-art approaches to the railway vehicle operation and maintenance, the principles of RAMS method [7, 45] was used to improve especially reliability, maintainability and availability of the investigated locomotive type with "hybrid design" (old vehicle with modernised propulsion system) by simpler and faster analysis of data from the diagnostic system of the diesel engine through. For processing of diagnostic data, statistical methods for assessment of the diesel engine reliability were used, similar as in the [16]. Outcome of the solution should improve the locomotive operation economy.

The expected service life (durability) of traction rail vehicles is about 30 years, while the service life of individual components/subsystems is usually not the same. The body and chassis generally last longer [44, 50], but e.g. the diesel engine or control systems have a shorter service life. Developments in the field of technology are advancing rapidly, and thus the vehicles are becoming technically obsolete and economically and ecologically disadvantageous. The operating and maintenance costs of such vehicles increase with age. One possibility is to replace them with new vehicles, which is very costly. Another possibility is to modernize them, which is a more acceptable and economically viable option for a large number of rolling stock operators, in particular in Central and Eastern Europe.

At present, apart from technological and economic factors, an environmental factor is gaining significance in restoring vehicle parts to fitness (regeneration) [26]. The use of regenerated parts reduces negative impact of production processes on the environment [9]. The positive effects are especially in saving energy and material that are not consumed for new products.

There are numerous examples of locomotive modernisation [22] or freight wagons [36]. One of them, ŽOS Vrútky, a Slovak company is active both in locomotives [12] and passenger wagons. Example of modernisation of a diesel locomotive is described in [4] where the proposed solution, an electronic rotations and power governor of diesel engine, was applied. By this solution a new optimal operational characteristic were realized. Efficiency of the modernisation has been assessed and supported by an LCC (Life Cycle Cost) analysis.

Sometimes the companies modify only specific components of the locomotive drive [2]. Similarly, a small change in maintenance technology is often a way to improve the reliability of vehicles, for example by adjusting the cleaning of the fuel system as mentioned in [14] or by maintaining the technical condition of the fuel system components [33]. For example, using powder details in various units of rolling stock proved to be more reliable, safe and economically profitable [27]. Usually, the benefits of modernization are in the improvement of economic and environmental parameters of rolling stock operation, for example by reducing a vehicle weight by using light materials [52].

In many European countries, including Slovakia, there are locomotives, which are obsolete and technically less suitable, are still in service on the railways. Therefore, the issues and tasks of modernization of diesel locomotives are very important. The aim of modernization is a positive change or affecting several important parameters of a locomotive at the same time. In the first place, these are operating parameters such as safety, reliability, energy efficiency, performance and much more. Another important factor of modernization are design improvements, which mainly result in less demanding maintenance, improvement of the overall care of the traction vehicle, simplification of operation, comfort and good ergonomic of a train driver and the like. Last but not least, the modernization also improves a number of environmental parameters of the vehicle [48], because modern construction elements and equipment applied in a modernized vehicle give priority to the maximum elimination of negative effects on the environment in the vicinity of railway lines from the operation of diesel traction vehicles.

The advantage of built-in diagnostics [8] is the interconnection of a large number of components, which provides the basis for a solid overview of the technical condition of most components as well as their functionality. A certain disadvantage of electronic diagnostics is the rather large number of error codes and the subsequent hierarchy of faults at several levels, which in part complicates the accurate identification of a specific problem. Diagnostics applied in the locomotive also facilitates the work of workers in repair/maintenance, in finding and identifying the specific cause of the failure. Another concrete application of diagnostic system used on diesel locomotives can be found in [1], which is used for checking the technical condition of some systems on the diesel Diesel locomotive, namely the electronic system for measuring, controlling and monitoring the speed and consumption of fuel as well as for control and signalling system.

As can be seen from the literature research, various researches and scholars use a variety of approaches to design an appropriate maintenance methodology. Some approaches are more focused on pure theoretical solutions to the maintenance problem, other are more practical oriented. However, each industry is specific and uses its methods. Different specifics are in industrial production and others has the transport sector. However, the suitability of the method used also affects the age, resp. date of manufacture of the technical equipment. In case of means of transport, e.g. in the automotive industry, where the requirements for operation efficiency and ecology change significantly every 5 years and the average age of cars in the EU is about 11 years, the development of maintenance methodologies is also more intense than, for example, railway vehicles, where the lifetime counts for decades. This, of course, corresponds to equipment of maintenance facilities with appropriate maintenance technology qualification of personnel. Due to the high investment in the purchase of new rolling stock, another specific feature of transport companies is that after the end of life of rolling stock they try to modernize vehicles, which leads to the creation of „hybrid“ designs, where a large part of the original, technically obsolete design is combined with a modern and economical propulsion unit, which is equipped with automatic control unit and fault evaluation. In such a case, it is necessary to choose an individual approach for the proposal of the maintenance methodology of the vehicle, taking into account all the factors mentioned. In our article we are dealing with such a "hybrid" case.

3. Subject of the study

The subject of the study is the locomotive series 757 (Fig. 1), which represents the latest project of the locomotive manufacturer for the modernization of diesel locomotives used for expressed trains haulage. The modernization was carried out from the original locomotive series 750 (T 478.0) or 754 (T 478.4), which were manufactured from the late sixties to the end of the seventies of the 20th century. The design change brought better operational and economic parameters as well as lower service costs [54].

The locomotive series 757 is a four-axle diesel-electric cabinet locomotive with alternate-direct current power transmission, total mass of 75.4 t and maximum operating speed 100 km/h. Its main utilisation is for medium heavy-duty rail track service on regional and state railway tracks with 1 435 mm gauge, in particular for passenger transport on non-electrified track of ŽSR (Slovak Railways). On the locomotive, an old diesel-electric generator unit was replaced by a new unit, composed of Caterpillar diesel engine, model 3512CHD with installed power of 1550 kW at 1800 rpm (188.5 rad/sec), traction alternator Siemens 1FW2 631-6 and auxiliary alternator Siemens 1FW4 630-10.

Fig. 1 Locomotive series 757[54]

The electric equipment includes an electronic control system MORIS RV07 [32]. The function of the modular control system MORIS RV07 is to control and monitor the parameters of components of diesel-electric locomotive with the purpose of simplifying the driver's control and reliability of operation. It controls the diesel engine Caterpillar as well as the alternators Siemens. The communication is executed with regulator through electric controllers and display unit of the control system PIXY with diagonal 10´´ on both control posts. The electronic system enables also the diagnosis of electronics, which creates a new information database on technical condition of the vehicle and new information links for assessment of scope, extent and duration of corresponding maintenance task for a particular vehicle.

The modernisation of the locomotive brought improvements in environmental parameters of the locomotive (lower fuel consumption and emissions, lower noise emissions, higher power). However, more frequent failures of the sophisticated control system consisting of a larger number of control and management elements appeared.

The output data system is complex and difficult to use for a locomotive driver as well as maintenance workers. For this reason, a simplification (user friendliness) of the output data system has been developed within our work to make it easier to identify and understand the data. By this, faster and clearer fault identification is achieved, which shortens maintenance time and increases the availability of locomotive utilisation in operation. This brings a direct economic effect in reduced maintenance costs as well as higher dependability of rail transport. Data from diesel engine control system can be used also for statistical processing after longer time of diesel operation and reveal the faultiest parts of the diesel engine, thus contribute to reliability improvement.

3. Data records processing and analysis

The technical condition of a vehicle is characterized by a relatively large set of parameters. In general, a parameter is a measurable quantity that describes the technical, economic or operational properties of an object. The limit values of the parameters in the vehicle are usually a criterion for the satisfactory function of the object and exceeding them is a criterion for failure. Each object is characterized by parameters that determine its qualitative indicators, either in terms of ensuring its basic characteristics and functional accuracy, or in terms of the effectiveness of its work, impact on the environment, etc. e.g. in the case of vehicles these can be speed, power, energy consumption, loading capacity, safety or mechanical and strength characteristics, kinematic and dynamic parameters of the vehicle or its components [5].

For determining the limit states and calculating the indicators of the so-called parametric reliability, it is important to thoroughly classify failures and determine for which failures it makes sense and whether limit states can be determined. To evaluate parametric reliability, the failures are divided according to the nature of the origin and course of processes leading to the fault.

Fig. 2. Control system MORIS RV07 [32]

The diagnostic system installed in the locomotive series 757 creates a large number of files in which it records operating parameters (physical and numerical). It creates in total 15 types of files. The created files are stored in the internal memory in the MORIS RV07 control system (Fig. 2), in which it creates a loop of data files for 2 weeks period. After this time, the oldest files are automatically deleted and replaced with the newly created files. The memory capacity is 5 GB and the stored files for the mentioned 2 weeks have a size of approximately 2.5GB. Files are saved in .dbf and .log formats.

Within the work, a CAT file was selected from all 15 types of files, which contains the physical and operational quantities of the locomotive's internal combustion engine and stores them in 17 subsystems. An example of a „CAT" file generated by the diagnostics is shown in Fig. 3.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	DATE	TIME	REZIM	OT_MER	OT_ZEL	LOAD	TV1	TV2	VYF_P	VYF_L	TLAK_OLEJ	TLAK_PAL	TLAK_TUR	AKT_SPOTR	MOTOHOD	SPOTREBA	POCET_ERR
2	3.3.2016	09:00:08	0	1104	1111	14	83	38	329	325	436	372	10	51	4257	458214	0
3	3.3.2016	09:00:09	0	1108	1109	13	83	38	328	324	436	372	10	51	4257	458214	0
4	3.3.2016	09:00:12	0	1109	1111	14	83	38	328	324	436	372	12	51	4257	458214	0
5	3.3.2016	09:00:13	0	1108	1110	14	83	39	328	324	436	372	12	51	4257	458214	0
6	3.3.2016	09:00:15	0	1109	1111	14	83	39	328	324	436	372	10	51	4257	458214	0
7	3.3.2016	09:00:17	0	1109	1110	14	83	39	328	324	436	372	10	51	4257	458214	0
8	3.3.2016	09:00:19	0	1108	1110	14	83	39	329	323	436	372	10	51	4257	458214	0
9	3.3.2016	09:00:21	0	1108	1111	14	83	39	329	323	436	372	10	51	4257	458214	0
10	3.3.2016	09:00:23	0	1110	1111	13	83	39	328	323	436	372	8	51	4257	458214	0

Fig. 3. An example of a „CAT" file generated by the diagnostics

The „CAT" file currently contains several quantities being recorded by the system. Specifically, these are the quantities that are summarized in the Tab. 1. However, some of these values do not have sufficient explanatory value for use in mathematical modelling or in real use. For this reason, we have made several changes in the „CAT" file. In the new "CAT" file, fields have been deleted that have no or only insignificant value for assessing the condition of the diesel engine and some important were added.

The first change in the "CAT" file was the recalculation of diesel engine rotations (rpm), measured (OT-MER) and required (OT-ZEL), to their difference. The difference between these two values has better informative value for the correctness of the diesel engine operation.

The second change was the addition of the value "VYF_ROZ", which records the difference in exhaust gas temperatures on the left and right side of the diesel engine, as it is a 'V' type engine (the cylinders in the form 'V' shape). This change will bring better condition monitoring of the group of cylinders on both sides of the diesel engine (temperature difference can signalize the variations in Air-fuel ratio, failure on injector, failure of valve, etc).

Table 1. Quantities recorded by the diesel engine diagnostics in the "CAT" file

Used abbreviations	Explanations
DATE	Date of values recording
TIME	Time of values recording
REŽIM	Mode selected by the locomotive operator
OT_MER	rpm measured
OT_ZEL	rpm required
LOAD	Instantaneous relative thrust
TV1	Coolant temperature 1st circuit
TV2	Coolant temperature 2nd circuit
VYF_P	Exhaust gas temperature on the left
VYF_L	Exhaust gas temperature on the right
TLAK_OLEJ	Oil pressure
TLAK_PAL	Fuel pressure
TLAK-TUR	Turbocharger pressure
AKT-SPOTR	Instantaneous consumption
MOTOHOD	Hours
SPOTREBA	Instantaneous fuel volume in the tank
POCET_ERR	Number of error messages

Another change concerned the recording of the lubricating oil pressure. The pressure is currently sensed before and after the filter, but only one of the values is recorded. For this reason, it is not possible to objectively evaluate the instantaneous lubricating oil pressure in the lubrication system. By sensing the pressure before and after the filter and comparing them, we have full control over this system and we can evaluate its condition in a short time. Based on this, we added a recalculation of the oil pressure difference before and after the filter. This fact will give us the opportunity to monitor the correct operation of the oil pump, the condition of which will be evaluated by the pressure sensor before the filter and at the same time the condition of the lubricating oil filter, the condition of which is monitored by the pressure difference before and after the filter. The situation is similar with fuel pressure. The fuel pressure is monitored before and after the filter, but only one of these values is recorded in the „CAT" file. By adding of the second value and at the same time their difference in to the new "CAT" file, we get complete control over the operation of the fuel pump and the condition of the fuel filter.

The changes were implemented in the form of adding new recorded parameters and modifying the already existing values.

Table 2. Explanation of creation of limits (intervals)

<AA, BB, CC, DD, EE>	AA, EE - parameters indicating fault limit values
	BB, DD - parameters indicating changes in the system leading to a fault
	CC - parameters indicating that the monitored system is OK

Table 3 Parameter intervals and differences in the new „CAT" file

Parameter	Interval / Difference
OT_ROZ	Difference max 50 rpm
TV1	<70 - 83 - 95 - 100 - 105>
TV2	<30 - 40 - 50>
VYF_P	<100 - 400 - 550 - 650 - 702>
VYF_L	<100 - 400 - 550 - 650 - 702>
VYF_ROZ	Difference between P and L max 30 ° C
TLAK_OLEJ_ROZ	Difference max 150 kPa
TLAK_PAL_ROZ	Difference max 150 kPa
TLAK_TUR	<150 - 210 - 250>

To evaluate the state (condition) of the diesel engine, it was necessary to define the limits (intervals) of the values of the selected parameters. In Table 2 the principle of creation of limits are explained.

Some parameters contained five values in the interval, for some it was sufficient to express three values. All parameters from the new „CAT" table and their intervals or differences between the two values were processed into Table 3.

Appropriate selection of intervals and values of differences from the real operation of locomotive series 757 were specially consulted with experts from the operation and maintenance departments of locomotive depots.

The next step was to create the „CAT" file itself in the „xlsx" format, in which all the changes mentioned above are incorporated. The new „CAT" file contains 10 000 data for each recorded quantity, which in total is approximately 90,000 values. The fault conditions were artificially changed beyond the intervals of Table 3. It is also important to mention that the fault-free values of the individual locomotive systems were selected from the locomotive diagnostics in a state where the locomotive had the driving mode selected, which is the most frequently used and most important in operation. Thus, the locomotive was in motion under load and the engine rotations were higher than 1450 rpm (151.8 rad/sec). For other modes (e.g. idling and transient modes) it is necessary to specify other intervals of parameter values (e.g. turbocharger pressure is lower at idling and run-out mode than in driving mode and rotations above 1450 rpm (151.8 rad/sec).

In Fig. 4, a part of the new „CAT" file is shown. There are measured and recalculated values of the engine parameters, which are compared with the defined intervals and values of the limit states of the parameters characterizing the failure.

The next step in preparing the data for evaluating the state of the diesel engine was to clearly define the faults using zeros and ones. If the parameter value fell outside the predefined interval, the parameter was assigned the number 0 (fault). If the parameter fell within the interval, it was assigned the number 1 (operation without failure). The overall condition of the locomotive was evaluated based on the condition of the individual locomotive systems (the column marked "LOKO" in the Fig. 5). If only one locomotive system acquires the value 0, then also in the column for the total locomotive state will be 0. The value 1 for the total locomotive state will be only if all the locomotive systems acquire a value of 1 and thus only then the locomotive is operational (up-state).

The original recording of the data did not provide the graphical output. A suitable tool for closer identification of the condition is a graphical representation of the course of the monitored values. The graphic visualisation simplifies the identification of incorrect values. As an example, the course of the engine rotations (rpm) is given, where the desired values are shown in red and the actually measured in green (Fig. 6). In the figure, an example of rotation stabilisation after their change is presented.

K	L	M	N	O	P	Q	R	S	T	U
LOKO		OT_ROZ	TV1	TV2	VYF_P	VYF_L	VYF_ROZ	TLAK_OLEJ_ROZ	TLAK_PAL_ROZ	TLAK_TUR
0		0	0	0	1	0	0	1	0	1
0		1	0	1	1	0	0	1	1	0
0		1	1	0	1	0	0	0	1	0
0		1	0	0	0	1	0	0	1	0
0		0	0	0	1	1	1	0	1	1
0		1	1	0	0	1	0	0	0	1
0		0	0	0	1	0	0	1	0	0
0		1	0	0	1	1	0	1	1	0
0		0	0	0	0	0	0	1	0	0
0		1	0	1	0	1	0	1	1	0
1		1	1	1	1	1	1	1	1	1
1		1	1	1	1	1	1	1	1	1
1		1	1	1	1	1	1	1	1	1
1		1	1	1	1	1	1	1	1	1
1		1	1	1	1	1	1	1	1	1

Fig. 5. Evaluation of the overall condition of the locomotive based on the individual systems condition

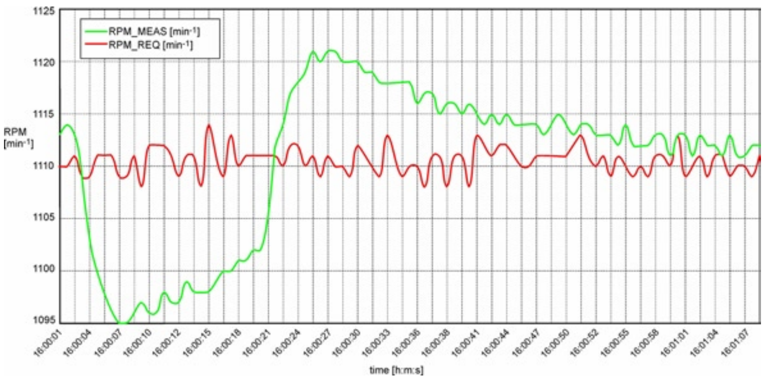


Fig. 6. Graphical representation of required and measured engine rotations

	A	B	C	D	E	F	G	H	I
1	OT_ROZ (ot/min)	TV1 (°C)	TV2 (°C)	VYF_P (°C)	VYF_L (°C)	VYF_ROZ (°C)	TLAK_OLEJ_ROZ (kPa)	TLAK_PAL_ROZ (kPa)	TLAK_TUR (kPa)
2	100	133	27	568	217	351	145	174	172
3	48	46	31	535	152	383	73	42	97
4	1	84	19	410	373	37	219	107	60
5	37	104	23	220	550	330	204	48	318
6	108	50	13	461	447	14	168	62	215
7	15	83	27	322	586	264	196	244	234
8	59	47	11	614	348	266	36	179	294
9	15	58	17	530	448	82	30	25	129
10	71	43	63	666	153	513	63	201	119
11	49	124	34	194	426	232	30	79	271
12	137	97	15	467	521	54	130	157	286

Fig. 4. Example of a modified set of measured data „CAT"

4. Improvement of the maintenance by new processing of data files – case example

Locomotive maintenance is often performed only on the basis of experience with its several years of operation. However, maintaining and diagnosing a locomotive in this way is very time-consuming and costly. It often happens that the locomotive returns to the depot after the maintenance, because the same failure reappeared despite the deployment of new components. Upon further inspection of the locomotive, it is determined that the failure might be caused by another component, which will then be replaced/repaired. Maintenance carried-out in this way is very inefficient in terms of time and finances.

Within the cooperation with the locomotive depot, which operates five locomotives series 757, an access to records of measurements of the operation of locomotives operated by the depot was provided. For the purposes of the work, the records on unplanned locomotive outages were separated from the records on all maintenance.

Records on unplanned outages of a particular locomotive 757. 016-1 (one of five in the depot) for the first four months of the year 2020 were processed (Table 4). On average, three to five unplanned outages per month occurred on this locomotive. The downtime ranged from one hour to repairs lasting several days. Interestingly, in March, faults occurred for three consecutive days, with one fault occurring repeatedly (heating failure) and the repair of the other (low insulation state of the excitation circuit) taking three days. This represents a very long downtime of the vehicle.

Every hour of an unplanned outage of the locomotive brings a loss to ZSSK – Slovak raiways, as the locomotive does not fulfil the tasks for which it was purchased. The average time of an unplanned locomotive outage is 4 days per month which is 48 days per year. The cost of unplanned outages of the locomotive must also include the cost of the work of the people performing the repair and diagnostics of the locomotive. The hourly work costs ranges from 20 to 28 Euros.

The vehicle maintenance time is divided into several parts. Some of them can be influenced and shortened (analysis of diagnostic files), some cannot be influenced directly (delivery of spare parts, repair by an external company). The proposed methodology reduces the time required for analyses of the files and work with them. Until now, the analysis of diagnostic files has been performed using a large number of files that contain a considerable amount of data and are often confusing. This fact affects and significantly extends the time needed to analyse the problem.

In the new method, after downloading the files, they are analysed and in this step the methodology can significantly reduce time. Only one ".log" file is analysed, which contains data for only one, specifically determined hour, using a record of the time of failure. The data is clearly arranged, it is possible to quickly and easily create a graph of parameters. The main advantage is the immediate display of data indicating the error and the display of the generated and signalised fault. This results in a faster return of the locomotive to operation, thus eliminating losses caused by the locomotive‘s downtime in maintenance. From experience in practical operation, the diagnostic process, which includes downloading files and analysing them, can currently take approximately 3 hours. Using the new method, this time can be reduced to an hour and a half, of which the analysis itself takes only half an hour.

5. Possibilities of further procedure in data analysis of large files in vehicle maintenance

When analysing and classifying data, e.g. from various electronic systems for monitoring the technical condition of vehicles to the planning and performance of maintenance from a statistical point of view, fuzzy files and logic become a valuable tool for modelling and processing inaccurate data, or for creating flexible techniques for handling accurate data. The so-called linguistic variables appear to be one of the promising ways of expressing values, described by quantitative or qualitative quantities. Qualitative quantities in many cases appear to be the result of the formalization of expert estimates. Each object or process is described by a group of indicators.

The use of the fuzzy method is applied in various fields, for example in the prediction of the reliability of structures, as mentioned in [13]. Prediction of structural performance is a complex problem because of the existence of randomness and fuzziness in engineering practice. In this area, reliability analyses have been performed using probabilistic methods. This work investigates reliability analysis of structure involving fuzziness and randomness. In particular, the safety state of the structure is defined by a fuzzy state variable, fuzzy random allowable interval, or fuzzy random generalized strength.

There are a number of methods for constructing a fuzzy set membership function based on expert estimates. Two groups of methods can be distinguished: direct and indirect. Direct methods assume that the expert immediately formulates rules according to which the value of the membership function characterizing the element is determined.

Table 4. An overview of faults occurred on the locomotive 757.016-1

Locomotive	Start of maintenance (date – time)	Finish of maintenance (date – time)	Fault description	Duration (hour:min)
757.016-1	16.01.2020 - 10:00	16.01.2020 - 12:00	Low insulation state of traction motor (TM)	2:00
757.016-1	03.02.2020 - 12:11	07.02.2020 - 17:00	4th TM faulty	100:49
757.016-1	09.02.2020 - 11:11	09.02.2020 - 15:00	Combustion engine in performance signals high crankcase pressure	3:49
757.016-1	25.02.2020 - 15:00	27.02.2020 - 17:00	Low insulation state of TM	50:00
757.016-1	10.03.2020 - 09:00	10.03.2020 - 18:00	Low insulation state of 3th TM	9:00
757.016-1	15.03.2020 - 07:00	15.03.2020 - 11:00	Non-functional train heating	4:00
757.016-1	16.03.2020 - 13:00	16.03.2020 - 14:00	Non-functional train heating	1:00
757.016-1	17.03.2020 - 11:11	20.03.2020 - 12:00	Low isolation state of the TM excitation circuit. Critically low oil pressure at start.	72:49
757.016-1	26.03.2020 - 04:48	26.03.2020 - 09:35	The locomotive without power	4:47
757.016-1	26.03.2020 - 17:28	27.03.2020 - 08:20	Defective primary circuit cooling inverter	14:52
757.016-1	07.04.2020 - 16:45	08.04.2020 - 17:00	Insufficient power of train heating	24:15
757.016-1	11.04.2020 - 04:00	12.04.2020 - 12:00	Fault 88.04 - power supply for NOV sensors, source of traction current sensor. The brake rod lock is missing on the 4th axis on the left.	32:00
757.016-1	23.04.2020 - 12:00	23.04.2020 - 17:00	Charging circuit	5:00

Indirect methods for calculating the values of the membership function are used when there are no elementary measurable properties.

The decision-making process is the most important moment in the management of various objects or processes, to which can be assigned the process of planning and implementation of vehicle maintenance based on the collection of large amounts of data from electronic systems monitoring their actual technical condition during operation. An essential component of this process is the selection of a decision from a set of acceptable alternatives. In many cases, the analysis of input situations as well as the selection of the best decision is carried out by comparison with decisions that have been made in the past, e.g. established maintenance system of the relevant vehicle. At the same time, it is necessary to minimize the costs of analysing input situations by determining the sequence of the most important indicators. The solution of the problem of recognizing the situation in decision-making is expressed in the form of analytical expressions or in the form of so-called Decision Tree.

The decision tree is created on the basis of a decision table describing N input situations (data and data on the technical condition of the vehicle, measuring and other samples). Each example is made up of the values of the input and output attributes, which for the maintenance of the vehicle means preparation, and the process of managing and performing the maintenance itself.

Fig. 7 shows the sequence of steps and operations required to make a decision tree.

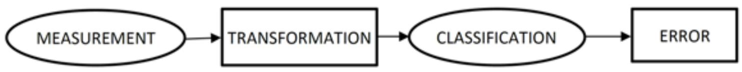


Fig. 7. Decision tree creation sequence scheme

The fuzzy logic method is suitable for analysing data from the diagnostics of complex systems [38]. It can be used also for diagnostics of locomotive series 757 due to the ability to process large amounts of data generated by the vehicle control system. The compilation of a decision tree for a diesel-electric locomotive series 757 in order to identify data from the electronics of the locomotive control systems focused on the maintenance of this locomotive will be the aim of another solution.

6. Conclusions

Safety, reliability, maintainability and operability are nowadays highly monitored parameters of locomotives performing line service on ŽSR lines. Quick maintenance is one of the most important indicators of the efficiency of locomotive operation. Therefore, any reduction in maintenance time is beneficial, as presented in the case of the locomotive series 757. Suggestions for improving the evaluation of the locomotive state based on diagnostics are beneficial for their practical use.

The specific changes are summarized in the following points:

1) creation of intervals (limits) for selected parameters of the internal combustion engine of the locomotive series 757, on the basis of which the algorithmic calculation evaluates the occurrence of the ICE failure,

2) the developed methodology (method) can display the parameters of the locomotive series 757 faster, easier and clearer and is able to immediately recognize the locomotive fault or faults based on the diagnostically created „.log" file,

3) the methodology enables the sorting of the monitored locomotive parameters and the display of selected parameters in a graph, which has an important benefit in practice in terms of the possibility of comparing two interdependent parameters of the locomotive,

4) creating a diagnostic report has benefits and advantages in terms of maintenance, namely in the area of better registration of interventions on the locomotive, control and registration of faults and the actual maintenance performances, maintenance rationalization and assessment of the results of the diagnostics itself,

5) based on long-term monitoring, the data can be more precisely statistically evaluated with the purpose of identification of individual failures and to define critical components of the propulsion unit.

Acknowledgements

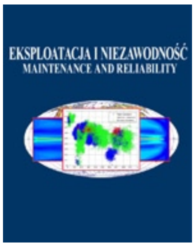
The paper was supported by the Scientific Grant Agency of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences in project KEGA, no. 018ŽU-4/2018: "Innovation of didactic approaches and content of subjects of technical diagnostic as a tool for enhancing the quality of professional knowledge for practice needs" and KEGA, no. 044ŽU-4/2019: Implementation of innovative elements in the education process within the study program Maintenance of means of transport".

References

1. Alexandrov M, Blanari I. Computerized diagnosis of electric diesel locomotive type 060 DA 2100 HP using the TATUNG TX-2000 WebPAd. IOP Conf. Series: Materials Science and Engineering 2018; 444 (6), https://doi.org/10.1088/1757-899X/444/6/062014.
2. Antipin DY, Vorobiev VI, Izmerov OV. Possibilities of modernization of wheel motor blocks of locomotives. IOP Conf. Series: Earth and Environmental Science 2019; 378, https://doi.org/10.1088/1755-1315/378/1/012004.
3. Antoni M. Formal validation method and tools for French computerized railway interlocking systems. International Journal of Railway 2009; 2(3): 99-106, https://doi.org/10.1109/ICCIE.2009.5223968.
4. Babel M, Szkoda M. Diesel locomotive efficiency and reliability improvement as a result of power unit load control system modernisation. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(1): 38-49, https://doi.org/10.17531/ein.2016.1.6.
5. Ben-Daya M, Duffuaa SO, Raouf A, Knezevic J, Ait-Kadi D. Handbook of Maintenance Management and Engineering. Springer 2009; 737, https://doi.org/10.1007/978-1-84882-472-0.
6. Berrade MD, Scarf PA, Cavalcante CAV, Dwight RA. Imperfect inspection and replacement of a system with a defective state: A cost and reliability analysis. Reliability Engineering and System Safety 2013; 120: 80-87, https://doi.org/10.1016/j.ress.2013.02.024.
7. Birolini A. Reliability engineering: Theory and Practice. Heidelberg: Springer, 2007.
8. Bosso N, Gugliotta A, Zampieri N. Design and testing of an innovative monitoring system for railway vehicles. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 2018; 232(2): 445-460, https://doi.org/10.1177/0954409716675005.
9. De Souza V, Borsato M, Bloemhof-Ruwaard J. Designing eco-effective reverse logistics networks. Transdisciplinary Engineering: Crossing Boundaries 2016; 4: 851-860.
10. Faulin J, Juan Perez AA, Martorell Alsina S S, Ramirez-Marquez J E (Eds.). Simulation Methods for Reliability and Availability of Complex

Systems. Springer 2010, https://doi.org/10.1007/978-1-84882-213-9.
11. Garmabaki AHS, Ahmadi A, Ahmadi M. Maintenance optimization using multi-attribute utility theory. Current Trends in Reliability, Availability, Maintainability and Safety 2016; 13-25, https://doi.org/10.1007/ 978-3-319-23597-4_2.
12. Hrinak J, Barta D. Repair and partial modernization of electric locomotives 441 Series of "Serbija Voz" In the company ZOS Vrutky as, XXIV. International conference Current problems in Rail Vehicles 2019; part I, 231-240.
13 Huang ZH, Zhang X, Liu Y, Meng D, Wang Z. Enhanced sequential optimization and reliability assessment for reliability-based design optimization. Journal of Mechanical Science and Technology 2012; 26 (7): 2039-2043, https://doi.org/10.1007/s12206-012-0511-7.
14. Kagramanian A, Stankevich P, Aulin D, Basov A. Efficiency improvement of locomotive-type diesel engine operation due to introduction of resource-saving technologies for cleaning diesel and diesel locomotive systems. Procedia Computer Science 2019; 149: 264-273, https://doi.org/10.1016/ j.procs.2019.01.133.
15. Kang R, Zhang QY, Zeng ZG, Zio E, Li XY. Measuring reliability under epistemic uncertainty: Review on non-probabilistic reliability metrics. Chinese Journal of Aeronautics 2016; 29(3):571-9, https://doi.org/10.1016/j.cja.2016.04.004.
16. Kazantseva NK, Kharlamov EP, Tkachuk GA, Kazantseva TV. Assessment of the reliability of the locomotive based on statistical methods of quality management. IOP Conf. Ser.: Mater. Sci. Eng. 2019; 666, https://doi.org/10.1088/1757-899X/666/1/012054.
17. Khudonogov A M, Khudonogov I A, Dulskiy E Y, Ivanov P Y, Lobytsin I O, Khamnaeva A A. Reliability analysis of power equipment of traction rolling stock within the Eastern region. IOP Conf. Ser.: Mater. Sci. Eng. 2020; 760, https://doi.org/10.1088/1757-899X/760/1/012018.
18. Konowrocki R, Chojnacki A. Analysis of rail vehicles' operational reliability in the aspect of safety against derailment based on various methods of determining the assessment criterion. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 73-85, https://doi.org/10.17531/ein.2020.1.9.
19. Kumar G, Jain V, Gandhi O. P. Availability analysis of mechanical systems with condition-based maintenance using semi-Markov and evaluation of optimal condition monitoring interval. Journal of Industrial Engineering International 2018; 14: 119-131, https://doi.org/10.1007/s40092-017-0212-z.
20. Lazarev A A, Grishin E M, Galakhov S A, Tarasov, G V. Algorithms for locomotives maintenance schedule. IFAC - PapersOnLine 2019; 52(13): 951-956, https://doi.org/10.1016/j.ifacol.2019.11.317.
21. Lin J, Pulido J, Asplund M. Reliability analysis for preventive maintenance based on classical and Bayesian semi-parametric degradation approaches using locomotive wheel-sets as a case study. Reliability Engineering and System Safety 2015; 134: 143-156, https://doi.org/10.1016/j.ress.2014.10.011.
22. Liudvinavicius L, Dailydka S. The Aspects of Modernization of Diesel-Electric Locomotives and Platform for Transportation of Railway Switches in Lithuanian Railways. Studies in Systems Decision and Control 2017; 87: 135-216, https://doi.org/10.1007/978-3-319-51502-1_4.
23. Macián V, Tormos B, Herrero J. Maintenance management balanced scorecard approach for urban transport fleets. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 226-236, https://doi.org/10.17531/ein.2019.2.6.
24. Macián V, Tormos B, Ruiz S, Riechi J. Urban bus fleet maintenance costs: comparative analysis between diesel vs CNG fuelled vehicles. EuroMaintenance 2014; ISBN 978-952-67981-3-4.
25. Meier-Hirmer C, Riboulet G, Sourget F, et al. Maintenance optimization for a system with a gamma deterioration process and intervention delay: application to track maintenance. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2009, 223(3): 189-198, https://doi.org/10.1243/1748006XJRR234.
26. Merkisz J, Jacyna M, Merkisz-Guranowska A, Pielecha J. Exhaust emissions from modes of transport under actual traffic conditions. WIT Transactions on Ecology and the Environment 2014; 190: 1139-1150, https://doi.org/10.2495/EQ141062.
27. Mironovs V, Stankevich P, Eiduks J. Tendencies in the field of application of metal-powder materials for repair and maintenance of railway equipment. Procedia Computer Science 2019; 149: 344-348, https://doi.org/10.1016/j.procs.2019.01.146.
28. Młynarski S, Pilch R, Smolnik M, Szybka J, Wiązania G. A concept of reliability assessment simulation model using systems structural decomposition. Journal of KONBiN 2018; 46: 51-74, https://doi.org/10.2478/jok-2018-0023.
29. Młynarski S. Evolution of machine reliability and life and economics of operational use. Management and Production Engineering Review 2016; 7(4): 76-85, https://doi.org/10.1515/mper-2016-0038.
30. Młynarski S, Pilch R, Smolnik M, Szybka J, Wiązania G. Formation of koon Systems Reliability Estimated with Analytical and Simulation Calculation Methods. Journal of KONBiN 2017; 42: 255-272, https://doi.org/10.1515/jok-2017-0028.
31. Nachlas J A. Reliability engineering. Probabilistic models and maintenance methods. Boca Raton: CRC Press Taylor & Francis Group 2017, https://doi.org/10.1201/9781315307596.
32. NES Nová Dubnica s.r.o. RV series power controllers. [Online], 2010; (Cited: 15.3.2019). Available at: http://www.nes.sk/documents/catalogProduct/original/92/MORIS_RV_07.pdf.
33. Osipowicz T, Abramek K F, Matuszak Z, Jaśkiewicz M, Ludwinek K, Poliak M. The Analysis of Technical Condition Common Rail Fuel System Components. 11th International Science and Technical Conference Automotive Safety 2018; 1-8, https://doi.org/10.1109/AUTOSAFE.2018.8373304.
34. Osyaev A, Rauba A, Kosarev A. Algorithm of the target function of time between overhauls taking into account gradual and sudden failure of the locomotive equipment. TransSiberia 2018, MATEC Web of Conferences 239: 01029, 2018, https://doi.org/10.1051/matecconf/201823901029.
35. Pennacchi P, Chatterton S, Vania A, Xu L. Diagnostics of Bearings in Rolling Stocks: Results of Long Lasting Tests for a Regional Train Locomotive. IFToMM 2019; MMS 61: 321-335, https://doi.org/10.1007/978-3-319-99268-6_23.
36. Placzek M, Wrobel A, Buchacz A A. Concept of technology for freight wagons modernization. IOP Conference Series-Materials Science and Engineering 2016; 161(1), https://doi.org/ 10.1088/1757-899X/161/1/012107.
37. Ramezani S, Moini A, Riahi M. Prognostics and health management in machinery: A review of methodologies for RUL prediction and roadmap. International Journal of Industrial Engineering and Management Science 2019; 6(1): 38-61.
38. Roumila Z, Rekioua D, Rekioua T. Energy management based fuzzy logic controller of hybrid system wind/photovoltaic/diesel with storage battery. International Journal of Hydrogen Energy 2017; 42(30): 19525-19535, https://doi.org/10.1016/j.ijhydene.2017.06.006.
39. Sangiorgio V, Mangini A M, Precchiazzi I. A new index to evaluate the safety performance level of railway transportation systems. Safety Science 2020; 131, https://doi.org/10.1016/j.ssci.2020.104921.
40. Sankararaman S, Goebel K. Uncertainty in prognostics and systems health management. International Journal of Prognostics & Health

Management 2015; 6: 1-11.
41. Selech J, Andrzejczak K. An aggregate criterion for selecting a distribution for times to failure of components of rail vehicles. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 102-111, https://doi.org/10.17531/ein.2020.1.12.
42. Sitarz M, Chruzik K. An approach to the legal requirements regarding railway transport safety monitoring in the european union. Transport 2019; 34(2): 163-174, https://doi.org/10.3846/transport.2019.8528.
43. Song H, Schnieder E. Evaluating Fault Tree by means of Colored Petri nets to analyze the railway system dependability. Safety Science 2018;110 Part A: 313-323, https://doi.org/10.1016/j.ssci.2018.08.017.
44. Stastniak P, Moravcik M, Baran P, Smetanka L. Computer aided structural analysis of newly developed railway bogie frame. MATEC Web of Conferences 157: 02051, 2018; https://doi.org/10.1051/matecconf/201815702051.
45. Szkoda M, Kaczor G, Satora M. Methodology of building the strategy of maintenance of rail vehicles with the use of the rams analysis in the area of safety. Journal of KONBiN 2019; 49(2), https://doi.org/10.2478/jok-2019-0033.
46. Szkoda M, Satora M. Change in the maintenance strategy as a method of improving the efficiency of the process of operation of railway means of transport. BulTrans-2018, MATEC Web of Conferences 234: 05001, 2018, https://doi.org/10.1051/matecconf/20182340 01.
47. Szkoda M, Satora M, Konieczek Z. Effectiveness assessment of diesel locomotives operation with the use of mobile maintenance points. Archives of Transport 2020; 54(2), https://doi.org/10.5604/01.3001.0014.2622.
48. Tułecki A, Szkoda M. Ecology, energy efficiency and resource efficiency as the objectives of rail vehicles renewal. Transportation Research Procedia 2017; 25: 386-406, https://doi.org/10.1016/j.trpro.2017.05.416.
49. Werbińska-Wojciechowska S. Preventive Maintenance Models for Technical Systems. Technical System Maintenance. Springer Series in Reliability Engineering 2019, https://doi.org/10.1007/978-3-030-10788-8.
50. Xiu R, Spiryagin M, Wu Q, Yang S, Liu Y. Fatigue life assessment methods for railway vehicle bogie frame. Engineering Failure Analysis 2020; 116, https://doi.org/10.1016/j.engfailanal.2020.104725.
51. Xu L, Pennacchi P, Chatterton S. A New Method for the Estimation of Bearing Health State and Remaining Useful Life Based on the Moving Average Cross-Correlation of Power Spectral Density. Mechanical Systems and Signal Processing 2020; 139, https://doi.org/10.1016/j.ymssp.2020.106617.
52. Zaripov R, Gavrilovs P. Research Opportunities to Improve Technical and Economic Performance of Freight Car through the Introduction of Lightweight Materials in their Construction, Procedia Engineering 2016; 187: 22-29, https://doi.org/10.1016/j.proeng.2017.04.345.
53. Zhao Y X. On preventive maintenance policy of a critical reliability level for system subject to degradation. Reliability Engineering and System Safety 2003; 79: 301-308, https://doi.org/10.1016/S0951-8320(02)00201-6.
54. ZOSZV, Product sheets (online), 2018, (Cited: 15.3.2019). Available at: http://www.zoszv.sk/userfiles/image/produktove%20listy/757%20PL%20vsetky%20strany.pdf



Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Özcan E, Yumuşak R, Eren T. A novel approach to optimize the maintenance strategies: a case in the hydroelectric power plant. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 324–337, http://doi.org/10.17531/ein.2021.2.12.

A novel approach to optimize the maintenance strategies: a case in the hydroelectric power plant

Indexed by: Web of Science Group

Evrencan Özcan[a], Rabia Yumuşak[a], Tamer Eren[a,*]

[a]Industrial Engineering Department, Kırıkkale University, 71451, Kırıkkale, Turkey

Highlights

- The maintenance strategy optimization model, which takes into account the wear rate, has been proposed.
- AHP, COPRAS and integer programming methods are used.
- The problem is addressed in the system dimension.
- A feasible model has been proposed.

Abstract

Countries need to develop sustainable energy policies based on the principles of environmental sensitivity, reliability, efficiency, economy and uninterrupted service and to maintain their energy supply in order to increase their global competitiveness. In addition to this impact of sustainable energy supply on the global world, maintenance processes in power plants require high costs due to allocated time, materials and labor, and generation loss. Thus, the maintenance needs to be managed within a system. This makes analytical and feasible maintenance planning a necessity in power plants. In this context, this study focuses on maintenance strategy optimization which is the first phase of maintenance planning for one of the large-scale hydroelectric power plants with a direct effect on Turkey's energy supply security with its one fifth share in total generation. In this study, a new model is proposed for the maintenance strategy optimization problem considering the multi-objective and multi-criteria structure of hydroelectric power plants with hundreds of complex equipment and the direct effect of these equipment on uninterrupted and cost-effective electricity generation. In the model, two multi-criteria decision-making methods, AHP and COPRAS methods, are integrated with integer programming method and optimal maintenance strategies are obtained for 571 equipment.

Keywords

maintenance management, maintenance strategy optimization, multi-criteria decision making, AHP, COPRAS, integer programming.

Abbreviations: AHP, Analytic hierarchy process; COPRAS, Complex Proportional Assessment; ELECTRE, ELimination Et Choix Traduisant la REalité; GP, goal programming; IP, integer programming; MCDM, multi-criteria decision-making; HPP, hydroelectric power plant; MSO, maintenance strategy optimization; SAW, Simple Additive Weighting; TOPSIS, Technique for Order of Preference by Similarity to Ideal Solution; VIKOR, Visekriterijumska Optimizacija I Kompromisno Resenje.

1.Introduction

Maintenance is a set of activities to assess and maintain the capabilities of instruments or equipment. However, another task of maintenance is to restore machinery or equipment that has lost its function back to its previous state [37]. Maintenance and repair activities are of great importance for any industrial plant to achieve sustainable generation, because failures resulting from improper planning of maintenance lead to generation losses and costly maintenance and repair expenses with the halt of generation. In addition to maintenance costs, generation halts endanger supply security, rendering enterprises unable to compete in highly competitive markets. Moreover, poor management of the maintenance process can lead to businesses being eliminated from markets. Furthermore, maintenance is a costly process in terms of time, labor requirement and material [43]. As maintenance planning is important and costly, it is critical to determine optimal maintenance strategies to be applied to the machine or equipment. This is because wrong maintenance strategies applied will generate major obstacles in achieving sustainable generation. For example, it not only increases the likelihood of equipment failure but also leads to high maintenance costs and reduced product quality [26]. Considering many factors such as cost, security of supply and product quality, it can be concluded that determining the maintenance strategies to be applied to the equipment is an indispensable first stage of an effective and feasible maintenance planning. This is because all maintenance and repair activities are performed according to the selected maintenance strategies [52]. There are many maintenance strategies in the literature: reliability-based maintenance [64], condition-based maintenance [2], risk-based maintenance [48], preventive maintenance [32], predictive maintenance [39], corrective maintenance [60], and lastly

(*) Corresponding author.
E-mail addresses: E. Özcan - evrencan.ozcan@kku.edu.tr, R. Yumuşak - rabiayumusak95@gmail.com, T. Eren - teren@kku.edu.tr

revision maintenance [44]. Among these strategies, four maintenance strategies are applied in HPP, where this study is implemented:

Corrective Maintenance Strategy: This maintenance strategy allows failure to occur before maintenance is performed. Corrective maintenance is a failure-based maintenance that is performed after a corrective or when an obvious probability of failure is detected. The purpose of this maintenance is to return the system to the state where it can perform its required function in the minimum possible time. A primitive type of maintenance, corrective maintenance does not take into account the losses caused by malfunctions and failures [60].

Preventive (Periodical) Maintenance Strategy: It is carried out according to predetermined periods or foreseen criteria. It is done to prevent the deterioration of the functioning of a product or to reduce the possibility thereof [32]. This type of maintenance aims to increase the reliability and availability of equipment by minimizing the number of failures and eliminating the need for unplanned corrective maintenance [61].

Predictive Maintenance Strategy: The goal of predictive maintenance is to reduce downtime and maintenance costs on the premise of zero failure generation by monitoring the operating status of the equipment and predicting when an equipment failure can occur [39]. Through prediction, it provides maintenance planning for future potential failures before the failure occurs. Ideally, the maintenance program is optimized to minimize maintenance costs and achieve zero failure generation [40].

Revision Maintenance Strategy: This is the maintenance strategy that involves the implementation of positive changes in the design, operation method, operating conditions, installation, scheduling and maintenance methods of the relevant machine/equipment in order to achieve the functions expected from the machine/equipment at the highest level [44].

As mentioned above, the most critical phase of maintenance management is maintenance planning. The first and indispensable stage of maintenance planning is the selection of the appropriate maintenance strategy. This selection problem is a very complex problem due to the fact that the system units have many and different functions, obtaining the data reflecting the system is difficult, and it contains many quantitative and qualitative criteria [42]. To solve this problem, researchers presented their solutions by using different methods in different application areas. With the recognition of the importance of maintenance management, the interest in the problem of maintenance strategy selection has increased in the literature in recent years. Increasing interest led researchers to compile and review published studies and as a result, two literature reviews on this subject were published in 2015. One of these two reviews was written by Ding and Kamaruddin [16]. In this review, researchers explained the problem of maintenance strategy selection in detail and evaluated the studies in a broad perspective and classified them into three groups. The other review was conducted by Shafiee [52]. Unlike Ding and Kamaruddin's study, Shafiee limited the studies on the basis of the methods used and evaluated them from a different perspective. Because of the multi-criteria and multi-objective structure of the problem, multi-criteria decision making methods are among the most preferred methods. Shafiee [52] evaluated this situation in detail by examining the MCDM methods used for maintenance strategy selection. Among the multi-criteria decision making methods, the most commonly used methods for the maintenance strategy selection problem in the literature are AHP [25], ANP [34], TOPSIS [17], SAW [50], ELECTRE [58], and VIKOR [38]. Instead of finding a solution to the problem of selecting a maintenance strategy using only one multi-criteria decision making method, some researchers have solved the problem using a combination of multi-criteria decision making methods. By using a combination of different decision making methods, these researchers have provided a different perspective to the problem of maintenance strategy selection. This has been of interest to researchers, and as a result, new studies have been published using the combination of AHP-TOPSIS [46], ANP-TOPSIS [47], ANP-ELECTRE [14], FAHP-VIKOR [28], AHP-PROMETHEE [19], FAHP-CODAS [45], ANP-DEMATEL [1], and AHP- COPRAS [22] to solve the maintenance strategy selection problem. The analytical level of the solution increased with the combined use of multi-criteria decision making methods, but these methods were insufficient in systems where the problem size increased. At this point, researchers have solved the problem of maintenance strategy selection for multiple equipment by using GP, one of the multi-criteria decision making methods [4, 8, 24, 29]. At this stage, the maintenance strategy selection problem has been replaced by MSO. GP method can also be integrated with multi-criteria decision making methods. For example, Bertolini and Bevilacqua [8] have determined the optimal maintenance strategy for centrifugal pumps in an oil refinery with the integration of AHP and GP methods. GP method has taken its place in the literature as a solution method for multi-objective MSO for multiple equipment. However, presence of more than one goal in GP method increases the complexity the problem and generates problems in obtaining the optimal result. In addition, more than one goal brings out the need for more data. Because of these disadvantages, research has shifted to IP as an alternative to GP method [51]. For example, Braglia et al. [9] used the failure mode effect analysis and IP methods to determine the costs of each strategy and which maintenance implementation was applicable to each failure. In this study, IP method was used since a single-goal model aimed at minimizing generation downtime was established and an optimal solution was sought for a very complex MSO problem since the plant consisted of hundreds of equipment.

When the application areas of the MSO problem, are examined, there are studies in many sectors including transportation [25], automotive [35], textiles [55] and machining [62]. There are many studies in the energy sector in which this study is conducted [53]. Williams and Patelli [23] found the optimal maintenance strategy for the IEEE-24 RTS equipment in a HPP with the Monte Carlo Simulation. Özcan et al. [44] performed a multi-objective MSO for 9 critical equipment in a HPP using AHP-TOPSIS and GP methods. In another study, Özcan et al. [42] calculated the criticality levels of the equipment in a HPP with AHP-TOPSIS methods. They proposed a model aimed at cost minimization by using these calculations in IP method. As a result, Özcan et al. [42] obtained optimal maintenance strategies for the seven electrical equipment groups. In the present study, a mathematical model was proposed to determine the maintenance strategies to be applied to all electrical equipment in a HPP MSO was performed for a total of 571 equipment. The model included the four maintenance strategies described in detail above. In the solution methodology, AHP and COPRAS –two multi-criteria decision making methods– were used for the calculation of some parameters. These parameters were then used in the IP model to obtain optimal maintenance strategies for 571 equipment. Based on the results obtained, the contributions of the model to the literature are as follows:

- MSO problem was solved within the system at a power plant for the first time. For example, while only critical equipment has been identified in the power plant and solutions have been proposed for only these equipment in the literature [42, 44], a solution was obtained in this study for all electric equipment in the plant. In the proposed model, optimal maintenance strategies of 571 equipment have been determined. With this study, a model yielding optimal results for such a large problem has been proposed for the first time in the literature within the context of MSO.
- Since the problem is handled at the system level and the plant consists of units, the problem includes identical equipment. However, since these equipment are located in different units, they have been subjected to different generation and maintenance processes. These generation and maintenance activities have generated wear differences between identical equipment. Due to failures caused by them, the resulting wear differences directly affect the maintenance strategy to be applied to identical equipment located in different units. In this study, the effect of wear was calculated by AHP which is one of the MCDM methods and reflected to the

model for the first time in the literature and real life consistency of the model was achieved.
- In the literature, single-goal models, usually involving cost minimization, or multi-goal models involving minimization of maintenance times, maintenance costs, downtime, etc., have been proposed in general. In this study, a model has been proposed to reflect the real life characteristics of the system by expressing many goals with a single goal -by minimizing the generation downtime of the system. In other words, the goal of minimizing generation stops generates a context including a set of objectives such as cost minimization, minimizing risk factors and reliability maximization.
- The integration of AHP-COPRAS-IP methods has been used for the MSO problem for the first time in the literature. In addition, the problem was removed from subjectivity by following a five step solution methodology. With the analytically obtained solution combination of decision problems within the scope of the study consisting of determination of equipment wear rates, determination of criticality levels of equipment for the plant and determination of the added value provided by each maintenance strategy to the plant, and optimal assignment of maintenance strategies to the equipment were achieved by taking into account the real life dynamics of the system.

In the second section of the study, the methods used and the reasons for using these methods are presented based on the advantages of the methods. In the third section, the application details of the study are presented, and in the fourth section and fifth section the results of the proposed model are evaluated and the study is completed by emphasizing the recommendations.

2. Methods

In this study, MSO problem of electrical equipment in a HPP is handled. First of all, the wear rate of nine units was calculated in order to reflect the differences of identical equipment to the model. Considering the multi-criteria structure of the problem, AHP method, which is frequently used in the literature and provides ease of use and flexibility in method integrations, was used for this calculation. In the second stage of the study, the added value of each maintenance strategy to the plant was calculated. At this stage, AHP method was used again because the multi-criteria nature and hierarchical structure of the problem. In the third stage of the problem, the criticality levels of the electrical equipment were determined. Although AHP-TOPSIS [44] integration is frequently used in the literature for this problem, AHP-COPRAS combination is used in this study. The equipment criticality levels need to be expressed over 100 in the mathematical model. Moreover, COPRAS method is more advantageous than TOPSIS method for the studies involving opposite criteria since the criteria are divided into two as useful and useless criteria and the algorithm is operated according to this separation [41]. Finally, the dynamics of the system are reflected in the model by using the parameters formed as a result of these three stages in the IP model. Details of the methods are provided further down in this section.

2.1. AHP

AHP is a method developed by Saaty and frequently used in many types of decision-making problems. This method gives the decision maker the opportunity to evaluate the criteria and alternatives in the decision-making process by analytically prioritizing them [49]. AHP is a widely used method in which ideas of groups are shared and the targets and alternatives are analyzed in order to obtain the best results. AHP method has been used in many fields of application including construction sector [15], health sector [11, 54], transport sector [10] and energy [56]. Furthermore, it has been preferred as a solution algorithm for many problems from efficiency assessment [56] to technology selection [5], from site selection [18] to maintenance planning [7, 13]. AHP method has been chosen as the solution method because it has the flexibility of integration with different methods such as reduction of subjectivity and linear programming and fuzzy logic [59].

Application steps of AHP are given below [49]:

Step 1: The purpose of the decision-maker is to include the criteria and alternatives that affect this purpose, and to determine the relationships between them and to compose a hierarchical structure.

Table 1. Saaty's preference scale [49]

Importance Values	Value Definitions
1	Equal importance of both factors
3	Factor 1 is more important than factor 2
5	Factor 1 is much more important than factor 2
7	Factor 1 has a very strong importance compared to factor 2
9	Factor 1 has an absolute superior importance to factor 2
2, 4, 6, 8	Intermediate values – when compromise is needed

Step 2: It is carried out by experts by comparing all criteria and alternatives according to their degree of importance. At this stage, the significance scale, which is developed by Saaty and given in Table 1, is used.

Step 3: Normalization process is done. This normalization is performed by dividing each value in each matrix by column totals (Eq. 1):

$$b_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} \qquad i, j = 1, 2, ..., n \tag{1}$$

Step 4: After normalization, the priority or weight vectors for the items compared in the hierarchy are calculated [Eq. 2]:

$$w_i = \sum_{i=1}^{n} b_{ij} \;/\; n \qquad i, j = 1, 2, ..., n \tag{2}$$

Step 5: The consistency rate (CR) is calculated. CR is calculated by applying equations Eq. 3, Eq. 4, Eq. 5, Eq. 6 respectively:

$$E_i = d_i \;/\; w_i \qquad i = 1, 2, \ldots, n \tag{3}$$

$$\lambda = \sum_{i=1}^{n} E_i \;/\; n \qquad i = 1, 2, \ldots, n \tag{4}$$

$$CI = (-n) / (n-1) \tag{5}$$

$$CR = CI \;/\; RI \tag{6}$$

Table 2. RI Values

n	1	2	3	4	5	6	8	9	10	11	12	13
RI	0	0	0,58	0,9	1,12	1,24	1,41	1,45	1,49	1,51	1,48	1,56

Table 2 is used to calculate Eq. 6. If *CR* less than 0.1 indicates that the application is consistent. Otherwise, the pairwise comparison matrices are revised, and the steps are repeated.

Step 6: In the analysis phase of AHP scores, the highest value alternative is chosen as the best alternative.

2.2. COPRAS

COPRAS, which is one of the MCDM methods, was developed in 1996 [63]. COPRAS can be used for maximum and minimum criteria values in multi-criteria evaluation. COPRAS method can be easily applied to problems involving complex criteria and numerous alternatives. Thanks to these features, it has been applied in many different fields in the literature. COPRAS method successfully solved different problems in different sectors from agriculture [3] to information sector, from investment evaluation [27] to supply chain management [12]. One of the most important features of COPRAS method is that it shows the degree of benefit of alternatives. It compares the evaluated alternatives with each other and expresses in percentage how good or bad the other alternatives are. In addition, it evaluates the criteria as useful and useless criteria and eliminates the need to make calculations on opposite criteria [63].

Application steps of COPRAS are given below [63]:

Step 1. The first step is to compose the decision matrix. Decision matrix (X) is formed as shown in Eq. 7. m is the number of alternatives and n is the number of criteria:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{7}$$

Step 2: In the second step, the decision matrix is normalized. Normalization process is carried out with the help of Eq. 8:

$$x_{ij}^{*} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} \quad j = 1,2,...,n \tag{8}$$

Step 3: The weighted normalized decision matrix is obtained by using the normalized decision matrix with the weight values of each evaluation criterion represented as *wj*. Normalized decision matrix expressed by *D* is formed with the help of Eq. 9:

$$D = d_{ij} = w_j * x_{ij}^{*} \quad i = 1,2,\ldots,m \quad j = 1,2,\ldots,n \tag{9}$$

Step 4: The sum of the values of the useful criteria in the weighted normalized decision matrix is shown as S_{i+}, for the useless criteria the sum is S_i-. Eq. 10 and Eq. 11 are used respectively for S_{i+} and S_i- calculations:

$$S_{i+} = \sum_{j=1}^{k} d_{ij} \quad j = 1,2,\ldots,k \tag{10}$$

$$S_{i-} = \sum_{j=k+1}^{n} d_{ij} \quad j = k+1, k+2,\ldots,n \tag{11}$$

Step 5. The relative importance value (Qi) is calculated using Eq. 12 for each alternative. The alternative with the highest Qi value means the best alternative:

$$Q_i = S_{i+} + \frac{\sum_{i=1}^{m} S_{i-}}{S_{i-} * \sum_{i=1}^{m} \frac{1}{S_{i-}}} \quad i = 1,2,\ldots,m \tag{12}$$

Step 6: In this step, the highest relative priority value is determined with the help of Eq. 13:

$$Q_{max} = max\{Q_i\} \quad i = 1,2,\ldots,m \tag{13}$$

Step 7: In the last step, the performance index (P_i) for each alternative is calculated using Eq. 14. The alternative with a P_i of 100 is considered as the best alternative. The order in which the alternatives should be preferred is obtained by ordering the performance index in descending order:

$$P_i = \frac{Q_i}{Q_{max}} * 100\% \quad i = 1,2,\ldots,m \tag{14}$$

2.3. Integer Programming (IP)

IP is the solution method in which some or all of the problem variables take integer values. Gomory suggested that by making small changes with the intersecting planes in the simplex algorithm, integer results could be obtained, and this led to an important breakthrough in IP [30]. After Gomory's study, different types of integer programming such as 0-1 and mixed IP came to the fore with various studies. The general form of the IP model is given below [57]:

$$\begin{aligned} & \text{Max (Min)} \quad z = g_O(x_1, x_2, ..., x_n) \\ & \text{St.} \\ & g_i(x_1, x_2, ..., x_n) \begin{Bmatrix} \le \\ = \\ \ge \end{Bmatrix} b_i, \qquad i \in M \equiv \{1,2,...,m\} \\ & x_j \ge 0, \qquad j \in N \equiv \{1,2,...,n\} \\ & x_j = \text{integer} \qquad j \in I \subseteq N \end{aligned} \tag{15}$$

IP has taken its place in the literature with effective results for different kinds of problems in many areas such as transportation [31], health [6], industry [33], and energy [20, 21].

3. Case study

The MSO which is an indispensable first phase of maintenance planning for one of the large-scale HPP with a direct effect on Turkey's energy supply security with its one fifth share in total generation is addressed in the study. Besides their share in energy generation, HPP are of great importance for environmentally friendly electricity generation since they are one of the renewable energy sources. Moreover, the most problematic phase in electricity generation is electricity transmission. The problems experienced in this phase especially reduce the output of the plant. For this reason, in the present study, all electrical equipment in a HPP are handled within the system. Optimal maintenance strategies are obtained for 571 equipment in total. These equipment include current transformers, voltage transformers, breakers, disconnectors, main power transformers, drive motors, auxiliary transformers, excitation transformers, slipring and carbon brushes, relays, transformer expansion tanks, bushings, generator rotors and generator stators and subcomponents of these equipment groups. it is formed. The proposed mathematical model serves to identify optimal maintenance strategies for 571 electrical equipment in the plant, increasing both efficiency and energy supply security. This study consists of four basic stages. Firstly, the wear rate of nine units was calculated by AHP method in order to reflect the differences of identical equipment to the model. In the second stage of the study, the added value provided by each maintenance strategy to the plant was obtained by AHP method. In the third stage, criticality levels of the equipment examined in the study were calculated for the plant. AHP-COPRAS integration, which is a multi-criteria decision making method, was used for this calculation. Plant experts were consulted to collect data during the implementation steps of the AHP method, which was used in three stages. The data were obtained with the help of 8 power plant experts (industrial, electrical, electrical-electronic and mechanical engineers) each of whom had 10 to 25 years of experience in operation and maintenance of HPP and by taking into account the real life operating rules of the HPP. Finally, optimal maintenance strategies were obtained by using the parameters calculated in these three stages in the IP model. The implementation steps of the new model proposed in this study are summarized in Figure 1.

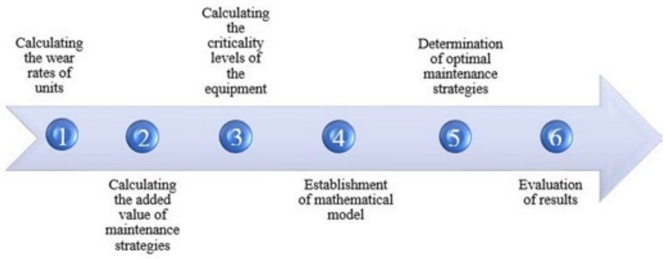


Fig. 1. Application steps

The power plant is composed of nine units. These units were commissioned at different times and different generation plans were implemented. This has caused wear differences between the units. One of the factors in determining the maintenance strategy is the wear rate of the equipment. This is because equipment with the same function and quality may require different maintenance practices as a result of different generation activities. Different maintenance practices require different maintenance strategies. In other words, wear rates affect MSO. In today's power plant operating conditions, it is not possible to calculate the wear rate of each equipment in HPP consisting of thousands of equipment, because of the difficulty in obtaining data and not being able to express completely different equipment with common parameters. In this study, calculating the wear rates of each unit was proposed as a solution.

3.1. Calculating the Wear Rates of Units

HPP are massive infrastructure investments. Therefore, it may not be possible to put all units into generation at the same time. The power plant discussed in this study consists of nine units. These units were activated at different times and different generation plans were implemented. This situation has caused differences in wear rate between the units. In fact, the model needs to be solved by taking into account the wear rate of each equipment. However, since the plant is composed of thousands of pieces of equipment and it is not possible to collect data for each equipment with common parameters in today's conditions, a MSO model that takes into account wear rates by calculating unit based wear rate is proposed for the first time in the literature. Three criteria were taken into account in the calculation of wear rate. These criteria are the date when the unit was commissioned, work time, and generation quantity. Wear rates for nine units were calculated according to these criteria. Considering the multi-criteria structure of the problem, AHP, which is the most used MCDM method in the literature by providing ease of solution to complex problems, was chosen. Expert opinions were used in the method. First, the criteria weights were generated. The steps of the AHP method described in Section 2.1 were applied at this stage. In the solution phase, a hierarchical structure was composed first. The hierarchical structure composed is given in Figure 2.

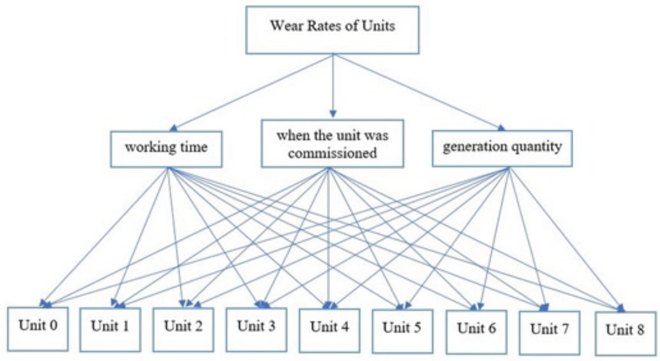


Fig. 2. Hierarchical structure

After the hierarchical structure was composed, it was passed to the stage where the weights of each criterion were determined. First, pairwise comparison matrices were composed. The pairwise comparison matrix composed is given in Table 3.

Table 3. The pairwise comparison matrix of the criteria

	When the unit was commissioned	Working time	Generation quantity
When the unit was commissioned	1	3	5
Working time	0.333	1	3
Generation quantity	0.200	0.333	1

For the calculation of criterion weights, the row averages of the values in the normalized decision matrix are taken. The weights formed after the process are given in Table 4.

As a result of the application, it is seen that the most important criterion is when the unit was commissioned with a weight of 0.63. This was followed by working time with a weight of 0.26. Finally, the weight of generation quantity was calculated as 0.11. The CR was 0.03.

Table 4. Criteria weights

Criteria	Weight
When the unit was commissioned	0.633
Working time	0.261
Generation quantity	0.106

Table 5. Weight vectors and CR values

When the unit was commissioned			Working time			Generation quantity		
Unit number	Weight vector	CR	Unit number	Weight vector	CR	Unit number	Weight vector	CR
U0	0.019	0.079	U0	0.017	0.067	U0	0.019	0.028
U1	0.028		U1	0.030		U1	0.030	
U2	0.028		U2	0.031		U2	0.047	
U3	0.251		U3	0.222		U3	0.303	
U4	0.056		U4	0.063		U4	0.047	
U5	0.251		U5	0.117		U5	0.212	
U6	0.056		U6	0.072		U6	0.077	
U7	0.071		U7	0.117		U7	0.110	
U8	0.242		U8	0.331		U8	0.156	

After the criterion weights were calculated, alternatives were evaluated for each criterion. Paired comparison matrices composed for each criterion are given in Appendix A. The results obtained when the steps given in Section 2.1 are applied in paired comparison matrices are given in Table 5.

Table 6. Wear rates of units

Unit number	AHP scores	Wear rates
U0	0.018	7.205
U1	0.028	11.13
U2	0.031	12.061
U3	0.249	97.187
U4	0.057	22.077
U5	0.212	82.778
U6	0.062	24.289
U7	0.087	33.979
U8	0.256	100

The wear rates of the nine units were calculated using the criterion weights obtained. By taking the ratio of the largest of the weights obtained by AHP to 100 and the wear rates were updated and re-expressed over 100. The results are given in Table 6. When the results are examined, it is seen that Unit 3, Unit 5 and Unit 8 are more worn than other units.

The benefits of each maintenance strategy to the plant are different. These differences are one of the main factors affecting the optimization of maintenance strategy. For this reason, in the second stage of the study, the added value provided to the plant by the four maintenance strategies discussed was measured.

3.2. Calculating the Added Value of Maintenance Strategies

The parameter that must be considered in problem solving for MSO is the added value of strategies provided to the whole system. This is because maintenance strategies have positive and negative effects reflected in the system in which they are applied. For example, the reduction of failures and increase in productivity as a result of the implementation of the maintenance strategy is a positive effect, while the cost items for the implementation of the strategy are a negative effect. For this reason, it is necessary to calculate the added value provided by the strategies to the plant and determine the maintenance strategy according to these values. In the present study, four maintenance strategies have been evaluated by taking into consideration the benefits, cost of maintenance process, duration and requirements for implementation of the strategy. This evaluation was made by AHP which is one of the multi-criteria decision making methods. The maintenance strategies implemented in the HPP detailed in Chapter 1 are summarized below.

Corrective Maintenance Strategy: Repair and/or maintenance activities carried out in the event that the machine/equipment is unable to perform the task expected of it, to ensure that the machine/equipment is capable of operating in line with its design specifications [44].

Preventive (Periodical) Maintenance Strategy: Maintenance activities carried out within a timetable for the machine/equipment to operate uninterruptedly and in line with the expected design specifications.

Predictive Maintenance Strategy: Maintenance activities which include monitoring of machine/equipment during operation by using modern measurement and signal-processing methods and taking necessary measures according to measurement results before failure occurs [43].

Revision Maintenance Strategy: It is a maintenance strategy which is done periodically (e.g. every 8000 hours or 5 years) to all critical equipment in the power plant units, which requires a long time (like 2 months) and in which the power plant unit downtime is mandatory [44].

The four maintenance strategies were evaluated under the criteria of benefit, cost, duration and requirements. First, a hierarchical structure was composed. The hierarchical structure composed is given in Figure 3.

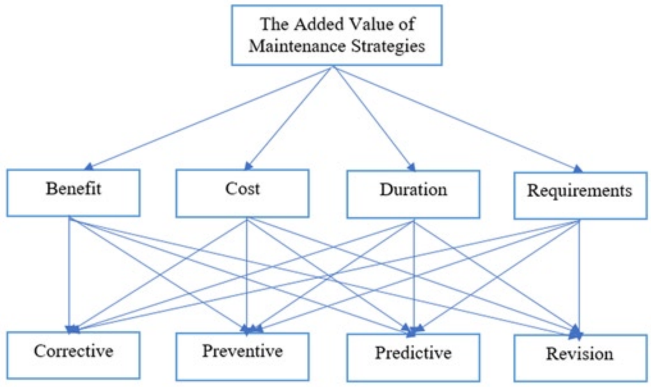


Fig. 3. Hierarchical structure

Table 7. The pairwise comparison matrix of the criteria

	Benefit	Cost	Duration	Requirements
Benefit	1	3	7	5
Cost	0.333	1	4	2
Duration	0.143	0.250	1	0.5
Requirements	0.200	0.500	2	1

Secondly, criteria weights were obtained by the AHP method. The pairwise comparison matrix of the criteria is given in Table 7.

Benchmark weights were found to be 0.579 for benefit, 0.233 for cost, 0.067 for duration, and 0.121 for requirements. When the criteria weights are evaluated, it is seen that the most important criterion is benefit. In the next step of the algorithm, the benefit values of the maintenance strategies were calculated by using the criterion weights obtained.

After the criterion weights were determined, the alternatives for each criterion were compared. Paired comparison matrices and CR values made in terms of criteria are given in Table 8.

It is seen that revision maintenance strategy provides the greatest added value. Corrective maintenance strategy is the maintenance strategy with the lowest added value. The results are given in Table 9.

Another factor affecting MSO is the criticality level of the equipment with respect to the power plant. In other words, it is a quantitative expression of the role of each equipment in electricity generation. In the third stage of the study, criticality levels of the equipment were calculated.

3.3. Calculating the Criticality Levels of the Equipment

The present study aims to assess and determine which maintenance strategies should be applied to 571 pieces of electrical equipment. In this problem, the maintenance strategy needs to be selected according to the equipment. Mathematical models should be used to obtain an optimal solution with a high analytical level not influenced by subjective judgments. In the mathematical model, qualitative data should be converted into quantitative data in order to reflect different aspects of the equipment. For this reason, the criticality levels of the equipment

Table 8. The pairwise comparison matrix of the criteria and CR values

Criteria	The pairwise comparison matrix of the criteria					CR
Benefit		Corrective	Revision	Predictive	Preventive	0.03
	Corrective	1	0.111	0.143	0.2	
	Revision	9	1	3	2	
	Predictive	7	0.33	1	0.5	
	Preventive	5	0.5	2	1	
Cost		Corrective	Revision	Predictive	Preventive	0.02
	Corrective	1	0.2	0.125	0.143	
	Revision	5	1	0.25	0.333	
	Predictive	8	4	1	2	
	Preventive	7	3	0.5	1	
Duration		Corrective	Revision	Predictive	Preventive	0.03
	Corrective	1	0.167	0.125	0.143	
	Revision	6	1	0.333	0.333	
	Predictive	8	3	1	0.5	
	Preventive	7	3	2	1	
Requirements		Corrective	Revision	Predictive	Preventive	0.06
	Corrective	1	0.167	0.111	0.333	
	Revision	6	1	0.25	4	
	Predictive	9	4	1	7	
	Preventive	3	0.25	0.143	1	

Table 9. Maintenance strategy added values

Maintenance strategy	Added values
Corrective	0.045
Revision	0.358
Predictive	0.325
Preventive	0.272

Table 10. The pairwise comparison matrix of the criteria

	C1	C2	C3	C4	C5	C6	C7
C1	1	0.200	0.250	0.125	2	0.500	0.500
C2	5	1	4	0.333	9	5	7
C3	4	0.250	1	0.200	4	3	2
C4	8	3	5	1	9	5	7
C5	0.500	0.111	0.250	0.111	1	2	0.333
C6	2	0.200	0.333	0.200	0.500	1	2
C7	2	0.143	0.500	0.143	3	0.500	1

Table 11. Criteria weights

Criteria	Code	Weight values
System backup	C1	0.042
Pre-maintenance conditions	C2	0.271
Failure period	C3	0.118
Possible results	C4	0.406
Availability of measuring equipment	C5	0.042
Processing time	C6	0.062
Fault detection	C7	0.059

for the plant should be determined. At this stage, the criteria affecting the criticality level were determined initially in accordance with the studies in the literature and expert opinions [42, 44]. Since the effect levels of these criteria are not equal, criterion weights should be determined. At this stage, weighting was performed with AHP, one of the multi-criteria decision making methods. The implementation steps given in Section 2.1 were carried out with the data obtained from the plant experts. The pairwise comparison matrix composed by the experts is given in Table 10.

The CR value was 0.089. The criteria and their weights are given in Table 11.

After determining the criterion weights, the necessary data to calculate the criticality levels of the equipment were collected. Data for seven criteria were collected for 571 pieces of equipment, but because the size of the data set is large, only the data for some pieces of equipment are given in Table 4. Using the data of 571 pieces of equipment, the criticality levels of the equipment were calculated by performing the COPRAS steps described in Section 2.2. Critical levels of some equipment are given in Table 12.

The aim of this study is to ensure that optimal maintenance strategies are assigned to 571 electrical equipment. In the first three sections, the parameters required for this purpose were obtained. Once

Table 12. Equipment data and critical levels

Equipment Name	C1	C2	C3	C4	C5	C6	C7	Critical level
6.3 KV Breakers	4	6	3	7	3	4	3	85.784
A Busbar Disconnector L1 Phase	1	7	5	10	3	4	3	100.000
Main Power Transformer Phase L1	1	7	3	10	3	4	3	95.716
Separator Motors L3 Phase	1	7	5	10	3	4	3	100.000
B Busbar Disconnector L2 Phase	1	7	5	10	3	4	3	100.000
Unpressurized Oil Tank Cooling Pump Drive Motor	4	7	3	1	1	2	1	51.344
BCT 19 (6.3 MVA) Transformer	4	1	5	7	3	2	3	70.006
Generator Group Breaker	4	1	5	9	1	2	3	75.897
Generator Rotor	1	7	5	10	3	4	3	100.000
Generator Stator	1	7	5	10	3	4	3	100.000
Internal Need Transformer	1	6	3	10	3	4	3	92.254

these steps were completed, the optimal solution of the problem was obtained with mathematical modeling.

3.4. Maintenance Strategy Optimization (MSO)

In the last stage of the study, MSO was performed for 571 pieces of equipment. An IP model was established with the values obtained in the first three stages. The objective of the mathematical model is to minimize generation downtime. In other words, the proposed new model aims to minimize generation downtime due to maintenance management in the plant as a result of optimal maintenance strategies to be implemented. Unlike multiple-goal models, this model optimizes only one parameter, but it has more than one goal. This is because it serves the most basic purpose of the maintenance process. This is to fulfill the role of all equipment in the system for the purpose of sustainable generation. Reducing generation downtime includes many goals such as minimizing costs, maximizing supply security, and risk minimization. For example, maximizing supply security depends on minimizing generation downtimes. Eliminating situations that may lead to generation downtimes will increase supply security. Or, minimizing generation downtime will keep failure risks to a minimum. As a result, determining the most appropriate maintenance strategies as described in the first section has a direct impact on the goal of sustainable generation. When this effect is taken into consideration, since the aim of the model is minimization of generation stops, it includes other goals as well.

The notations and decision variables used in the model are described below.

Notations:

i: Unit index (i=0,…,8)

j : Equipment index (j=1,…,68)

k : Maintenance strategy index(1=revision, 2=preventive, 3=predictive, 4=corrective)

T_{ijk} : ith unit, jth equipment production downtime when kth maintenance strategy is applied

D_{ijk} : ith unit, jth equipment kth maintenance strategy implementation time

C_{ijk}: ith unit, jth equipment kth maintenance strategy implementation cost (sum of labor and material cost)

Tc: Budget allocated for maintenance

$$X_{ijk} = \begin{cases} 1, & \text{if unit i assigned to j equipment in kth maintenance strategy} \\ 0, & \text{otherwise} \end{cases}$$

Td: Maintenance time (hours)

CR_{ij}: ith unit to critical level of jth equipment

Y_i: ith unit attrition rate

W_k: kth weight of maintenance strategy

Decision variables:

i= 0,…,m j= 1,…,n k= 1,…,l

Model formulation:

$$\text{Min } Z = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l} T_{ijk} * X_{ijk} \tag{1}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l} D_{ijk} * X_{ijk} \leq Td \tag{2}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l} C_{ijk} * X_{ijk} \leq \text{Tc} \tag{3}$$

$$\sum_{k=1}^{l} X_{ijk} \geq 1 \quad i=0,\ldots,m \quad j=1,\ldots,n \tag{4}$$

If $CR_{ij} \geq 85$

$$\sum_{k=1}^{l} W_k X_{ijk} \geq 0.85 \quad i=0,\ldots,m \quad j=1,\ldots,n \tag{5}$$

If CRij ≥ 70 V CRij < 51

$$\begin{aligned} &\sum_{k=1}^{l} W_k X_{ijk} \geq 0.70 \quad i=0,\ldots,m \quad j=1,\ldots,n \\ &\sum_{k=1}^{l} W_k X_{ijk} \leq 0.85 \quad i=0,\ldots,m \quad j=1,\ldots,n \end{aligned} \tag{6}$$

If CRij ≥ 51 V CRij < 70

$$\sum_{k=1}^{l} W_k X_{ijk} \geq 0.51 \quad i=0,\ldots,m \quad j=1,\ldots,n$$

$$\sum_{k=1}^{l} W_k X_{ijk} \leq 0.70 \quad i=0,\ldots,m \quad j=1,\ldots,n \tag{7}$$

$$X_{ij1}=0 \quad i=0,\ldots,m \quad j=1,\ldots,n$$

Else if $CR_{ij} < 51$

$$\begin{aligned} X_{ij1}&=0 \quad i=0,\ldots,m \quad j=1,\ldots,n \\ X_{ij2}&=0 \quad i=0,\ldots,m \quad j=1,\ldots,n \\ X_{ij3}&=0 \quad i=0,\ldots,m \quad j=1,\ldots,n \\ X_{ij4}&=1 \quad i=0,\ldots,m \quad j=1,\ldots,n \end{aligned} \tag{9}$$

İf $Y_i \geq 80 \land CR_{ij} \geq 70$

$$X_{ij2}=1 \quad i=0,\ldots,m \quad j=1,\ldots,n \tag{10}$$

Formulation of the mathematical model is given below. Eq. 1 represents the objective function of the model. It means minimization of generation downtime. Eq. 2 indicates that the actual maintenance period should be less than or equal to the assigned maintenance period. Eq. 3 means that the total maintenance cost should be less than or equal to the total budget allocated for maintenance. Eq. 4 means that at least one maintenance strategy must be assigned to each equipment. Eq. 5- Eq. 9 are the constraints that make the assignments by taking into account the criticality levels of the equipment. The sum of the added value obtained from the maintenance strategies to be implemented should be proportional to the criticality level of the equipment. The threshold values were determined according to the pre-maintenance conditions and possible results, which were determined as the two most important criteria in the calculation of the critical levels of the equipment described in Section 3.3. Eq. 10 stated that periodic maintenance should be performed if the wear rate of the unit i is greater than or equal to 80 and the criticality level of the equipment is greater than or equal to 70. This constraint is added for units with high wear because of the high possibility of equipment failure. The reason for limiting the level of criticality is because this maintenance cost must be borne for critical equipment.

4. Results and discussion

Maintenance is costly in terms of generation loss, time, labor and material requirements due to disruption of generation during the process, and is difficult to manage due to the inherent limitations of these components. In this context, MSO problem which is the indispensable first step of maintenance planning was discussed in this study. A HPP with nine units was investigated. In the HPP, which consists of thousands of equipment, electrical equipment was taken into consideration due to the major problems in the transmission of electricity. Optimal maintenance strategies were achieved for a total of 571 equipment. For these results, firstly the wear rate of nine units was calculated by AHP method in order to reveal the difference of wear between the units. Then, the benefit (added value) of the maintenance strategies to the plant was solved by AHP method. Afterwards, criticality levels of the studied equipment were solved by AHP-COPRAS integration. Three different parameters calculated were used in the 0-1 IP model. The objective of the mathematical model is minimization of generation stops. Minimizing generation downtime includes many goals such as minimizing costs, maximizing supply security, and minimizing risk. In this way, a single-goal model was used to reflect a multi-goal structure and a feasible model proposal was obtained. The model, whose canonical form is given in Section 3.4, has 2284 decision variables and 14 constraint sets. The model was solved by using ILOG CPLEX Studio IDE version 12.8. Optimal results were obtained in 1 second. As the number of equipment handled in this study was quite high, the results of all equipment could not be provided here. Several pieces of equipment with different wear rates and criticality levels were selected. The optimal maintenance strategies of these selected equipment are given in Table 13. All results of the model are given in Appendix B.

When the results of the mathematical model generated by IP method are evaluated, it is seen that if the criticality level of the equipment is greater than 85, all maintenance strategies except for corrective maintenance should be applied. This means that if the equipment is critical to the system, revision, periodic and predictive maintenance should be performed without waiting for equipment failure. This is because when these equipment fail; the unit shuts down and endangers energy supply security. For equipment with a criticality level of 70 to 85, revision, predictive and corrective maintenance strategies should be implemented. This is because this equipment does not cause generation downtime in case of failure, but generation resumes without backup. Operation without back-up (redundancy) poses the risk of generation downtime in case of any failure. In this case, major maintenance, which is revision maintenance, must be performed. In addition, equipment should be monitored continuously by predictive maintenance strategy. This monitoring will allow the equipment to be intervened before failure. In addition, if the equipment fails, corrective maintenance strategy should be applied. However, if the equipment is in one of the units with high wear rate, the probability of failure will be kept to a minimum by applying maintenance periodically instead of corrective maintenance. When equipment with a criticality level of 51 to 70 fails, the unit does not stop, but this may pose a problem in an emergency. For this reason, in order to prevent malfunctions, frequent periodical maintenance can be performed and monitoring the equipment regularly with predictive maintenance will be sufficient. Since equipment with a criticality level of less than 50 does not have any impact on the system –such as unit downtime or operation without backup–, only maintenance strategy that should be implemented is corrective maintenance. There are many academic studies in the literature to reduce maintenance costs and equipment failures in production facilities. Generally, a maintenance strategy that has to be implemented has been determined using a MCDM method for a single piece of equipment [26]. However, most production facilities, such as the hydroelectric power plant under consideration, consist of multiple intertwined equipment or sub-systems. This structure of the facility caused the necessity of determining the maintenance strategy within the system for the maintenance strategies determined by analytical methods to be applicable in the real manufacturing facilities. With this requirement, models providing MSO for more than one equipment have been proposed in the literature. Among these models, Bertolini and Bevilacqua [8], which consider the most equipment in the literature, discussed 10 centrifugal pumps. MSO for up to 14 equipment was performed for HPP [44]. In this study, a MSO was performed for all electrical equipment (571 equipment) in a hydroelectric power plant. Although the equipment features are the same, the wear and tear differences have occurred as a result of different maintenance and generation plans. Since these attrition differences are an important factor in determining the maintenance strategies to be applied, the attrition differences between the units are reflected in the proposed model. This approach has increased both the applicability of the optimal results to the real system and a MSO has been made by considering the attrition rates for the first time in the literature.

5. Conclusion

The main purpose of maintenance activities is to maximize the efficiency and effectiveness of production and increase reliability. This goal makes maintenance not an auxiliary process for production, mak-

Table 13. Optimal maintenance strategies of some equipment

Unit number	Equipment name	Criticality levels	Revision	Preventive	Predictive	Corrective
4	A busbar disconnector L3 phase	100	✓	✓	✓	
7	Separator motors L3 phase	100	✓	✓	✓	
5	B busbar disconnector L2 phase	100	✓	✓	✓	
1	Generator rotor	100	✓	✓	✓	
3	Generator stator	100	✓	✓	✓	
8	Main power transformer L1 phase	95.716	✓	✓	✓	
2	Warning transformer	95.716	✓	✓	✓	
5	Bearing Oil Pump Drives	92.62	✓	✓	✓	
1	Internal need transformer	92.254	✓	✓	✓	
8	Transformer bucholz relay	92.156	✓	✓	✓	
7	Transformer overcurrent relay	92.156	✓	✓	✓	
7	Transformer Expansion Tank	91.816	✓	✓	✓	
8	Transformer Expansion Tank	91.816	✓	✓	✓	
2	Slipring and carbon brushes	91.786	✓	✓	✓	
3	Slipring and carbon brushes	91.786	✓	✓	✓	
8	Transformer High Voltage Bushings	90.628	✓	✓	✓	
6	6.3 KV breakers	85.784	✓	✓	✓	
7	Servomotor pressure oil pumps drive motors	82.589	✓		✓	✓
8	Servomotor pressure oil pumps drive motors	82.589	✓	✓	✓	
1	Speed governor pressure oil pumps drive motors	78.305	✓		✓	✓
3	Speed governor pressure oil pumps drive motors	78.305	✓	✓	✓	
7	Speed regulator air compressors drive motors	78.305	✓		✓	✓
8	Speed regulator air compressors drive motors	78.305	✓	✓	✓	
0	Pump 1-2A drive motor	77.275	✓		✓	✓
0	Generator group breaker	75.897	✓		✓	✓
0	Deep well pump-1 drive motor	70.265	✓		✓	✓
0	BCT 19 (6.3 MVA) transformer	70.006	✓		✓	✓
0	BCT 22 (6.3 MVA) transformer	70.006	✓		✓	✓
0	High Pressure Air Compressor Drive Motors-a1	64.918		✓	✓	
0	220 V DC accumulators	59.9		✓	✓	
1	Cooler-1 fan-1	52.406		✓	✓	
6	Generator rotor lifting high pressure oil pump drive motor	51.923		✓	✓	
8	Generator rotor lifting high pressure oil pump drive motor	51.923		✓	✓	
6	Unpressurized oil tank cooling pump drive motor	51.344		✓	✓	
7	Leakage oil pump drive motor	35.069				✓
8	Exhaust fan drive motor	31.155				✓

ing it one of the basic processes for the production to reach a certain efficiency and efficiency target [36]. The indispensable and first step in managing this main process is MSO. In this context, in this study MSO problem is discussed in one of the large-scale HPP directly acting the Turkey's energy supply security.

This study includes many combinations of methods to increase the applicability of the problem to a real plant and to increase the level of analytics. In this study consisting of four basic phases, the attrition rate of nine units was calculated by AHP method in order to reflect the differences of identical equipment from each other to the model in the first phase. In the second stage of the study, the added value provided by each maintenance strategy to the power plant was again obtained through the AHP method. In the third stage, the criticality levels of the equipment discussed in terms of power plants were calculated. In this calculation, AHP-COPRAS integration was used. Finally, using the parameters calculated in these three stages in the IP model, optimal maintenance strategies were obtained for 571 equipment.

Although the proposed model deals with a HPP, various calculations have been made to reflect the dynamics of the system to the model. Although these calculations are made specific to the power plant under consideration, they can be adapted for other enterprises. Because in the model, the wear rate, the added value provided by the maintenance strategies to the system and the criticality levels of the equipment are calculated, and all these parameters are the factors that affect the selection of the maintenance strategy regardless of the system. However, the system under consideration should be analyzed in detail in order to adapt the proposed model to different businesses. The proposed model is flexible in terms of adapting the specific constraints of the system to the model.

Contrary to the literature, the power plant has been evaluated on a system basis and for the first time in the literature, an optimal solution of such a large problem has been proposed. In addition, due to the different attrition rates between units, a constraint was written according to the attrition rate and a solution for this situation was produced for the first time in the literature.

In the next stage of this study, mechanical equipment can be included with electrical equipment and the problem size can be increased. This will make it more difficult to obtain an optimal solution, therefore, intuitive approaches can be developed.

Acknowledgements

This study is supported by Kırıkkale University Scientific Research Projects Coordination Unit within the scope of scientific research project numbered 2019/087 and 2020/001.

Appendix A

Criteria	The pairwise comparison matrix of the criteria									
when the unit was commissioned		**U0**	**U1**	**U2**	**U3**	**U4**	**U5**	**U6**	**U7**	**U8**
	U0	1	0.333	0.333	0.111	0.2	0.111	0.2	0.111	0.2
	U1	3	1	1	0.125	0.333	0.125	0.333	0.2	0.125
	U2	3	1	1	0.125	0.333	0.125	0.333	0.333	0.125
	U3	9	8	8	1	7	1	7	7	1
	U4	5	3	3	0.14	1	0.143	1	1	0.143
	U5	9	8	8	1	7	1	7	7	1
	U6	5	3	3	0.14	1	0.14	1	1	0.143
	U7	9	5	3	0.14	1	0.14	1	1	0.143
	U8	5	8	8	1	7	1	7	7	1
working time		**U0**	**U1**	**U2**	**U3**	**U4**	**U5**	**U6**	**U7**	**U8**
	U0	1	0.333	0.333	0.125	0.2	0.167	0.2	0.167	0.111
	U1	3	1	1	0.167	0.2	0.2	0.25	0.2	0.143
	U2	3	1	1	0.167	0.333	0.2	0.25	0.2	0.143
	U3	8	6	6	1	5	4	5	4	0.333
	U4	5	5	3	0.2	1	0.25	0.5	0.25	0.167
	U5	6	5	5	0.25	4	1	3	1	0.2
	U6	5	4	4	0.2	2	0.333	1	0.333	0.167
	U7	6	5	5	0.25	4	1	3	1	0.2
	U8	9	7	7	3	6	5	6	5	1
generation quantity		**U0**	**U1**	**U2**	**U3**	**U4**	**U5**	**U6**	**U7**	**U8**
	U0	1	0.333	0.25	0.111	0.25	0.143	0.2	0.167	0.143
	U1	3	1	0.5	0.143	0.5	0.167	0.2	0.167	0.143
	U2	4	2	1	0.167	1	0.2	0.5	0.333	0.25
	U3	9	7	6	1	6	2	5	4	3
	U4	4	2	1	0.17	1	0.2	0.5	0.333	0.25
	U5	7	6	5	0.5	5	1	4	3	2
	U6	5	5	2	0.2	2	0.25	1	0.5	0.333
	U7	6	6	3	0.25	3	0.33	2	1	0.5
	U8	7	7	4	0.333	4	0.5	3	2	1

Appendix B

Equipment number	Revision	Preventive	Predictive	Corrective	Equipment number	Revision	Preventive	Predictive	Corrective	Equipment number	Revision	Preventive	Predictive	Corrective
1-19		✓		✓	193-195	✓	✓	✓		391				✓
20-21	✓	✓	✓		196-197		✓	✓		392	✓	✓	✓	
22-27		✓		✓	198				✓	393	✓		✓	

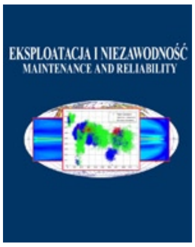
28-49	✓	✓	✓		199	✓	✓	✓		394	✓	✓	✓	
51				✓	200-217		✓	✓		395-396	✓		✓	
52	✓	✓	✓		218-240	✓	✓	✓		397-399	✓	✓	✓	
53	✓		✓		242-253	✓	✓	✓		400-401	✓		✓	
54	✓	✓	✓		254	✓		✓		402				✓
55-56	✓		✓		255				✓	403	✓	✓	✓	
57-59	✓	✓	✓		256	✓	✓	✓		404-421	✓		✓	
60-61	✓		✓		257	✓		✓		422-444	✓	✓	✓	
62				✓	258	✓	✓	✓		445	✓		✓	
63	✓	✓	✓		259-260	✓		✓		446-457	✓	✓	✓	
64-81	✓		✓		261-263	✓	✓	✓		458	✓		✓	
82-104	✓	✓	✓		264-265	✓		✓		459				✓
105	✓		✓		266				✓	460	✓	✓	✓	
106-117	✓	✓	✓		267	✓	✓	✓		461	✓		✓	
118	✓		✓		268-285	✓		✓		462	✓	✓	✓	
119				✓	286-308	✓	✓	✓		463-464	✓		✓	
120	✓	✓	✓		309		✓	✓		465-467	✓	✓	✓	
121	✓		✓		310-321	✓	✓	✓		468-469	✓		✓	
122	✓	✓	✓		322		✓	✓		470				✓
123-124	✓		✓		323				✓	471	✓	✓	✓	
125-127	✓	✓	✓		324	✓	✓	✓		472-489	✓		✓	
128-129	✓		✓		325		✓	✓		490-526	✓	✓	✓	
130				✓	326	✓	✓	✓		527				✓
131	✓	✓	✓		327-328		✓	✓		528	✓	✓	✓	
132-149	✓		✓		329-331	✓	✓	✓		529		✓	✓	
150-172	✓	✓	✓		332-333		✓	✓		530	✓	✓	✓	
173		✓	✓		334				✓	531-532		✓	✓	
174-185	✓	✓	✓		335	✓	✓	✓		533-535	✓	✓	✓	
186		✓	✓		336-353		✓	✓		536-537		✓	✓	
187				✓	354-376	✓	✓	✓		538				✓
188	✓	✓	✓		377	✓		✓		539	✓	✓	✓	
189		✓	✓		378-389	✓	✓	✓		540-557		✓	✓	
190	✓	✓	✓		390	✓		✓		558-571	✓	✓	✓	
191-192		✓	✓											

References

1. Aghaie M, Fazli S. An improved MCDM method for maintenance approach selection: A case study of auto industry. Management Science Letters 2012; 2(1): 137-146, https://doi.org/10.5267/j.msl.2011.09.012.
2. Alaswad S, Xiang Y. A review on condition-based maintenance optimization models for stochastically deteriorating system. Reliability Engineering and System Safety 2017; 157: 54-63, https://doi.org/10.1016/j.ress.2016.08.009.
3. Arabameri A, Yamani M, Pradhan B et al. Novel ensembles of COPRAS multi-criteria decision-making with logistic regression, boosted regression tree, and random forest for spatial prediction of gully erosion susceptibility. Science of The Total Environment 2019; 688: 903-916, https://doi.org/10.1016/j.scitotenv.2019.06.205.
4. Arunraj N S, Maiti J. Risk-based maintenance policy selection using AHP and goal programming. Safety science 2010; 48(2): 238-247, https://doi.org/10.1016/j.ssci.2009.09.005.
5. Bagassi S, De Crescenzio F, Piastra S. Augmented reality technology selection based on integrated QFD-AHP model. International Journal on Interactive Design and Manufacturing (IJIDeM) 2020; 14(1): 285-294, https://doi.org/10.1007/s12008-019-00583-6.
6. Bastos L S L, Marchesi J F, Hamacher S, Fleck J L. A mixed integer programming approach to the patient admission scheduling problem. European Journal of Operational Research 2019; 273(3): 831-840, https://doi.org/10.1016/j.ejor.2018.09.003.
7. Benítez P, Rocha E, Varum H, Rodrigues F. A dynamic multi-criteria decision-making model for the maintenance planning of reinforced concrete structures. Journal of Building Engineering 2020; 27: 100971, https://doi.org/10.1016/j.jobe.2019.100971.

8. Bertolini M, Bevilacqua M. A combined goal programming-AHP approach to maintenance selection problem. Reliability Engineering & System Safety 2006; 91(7): 839-848, https://doi.org/10.1016/j.ress.2005.08.006.
9. Braglia M, Castellano D, Frosolini M. An integer linear programming approach to maintenance strategies selection. International Journal of Quality & Reliability Management 2013; 30(9): 991-1016, https://doi.org/10.1108/IJQRM-05-2012-0059.
10. Bureika G, Bekintis G, Liudvinavičius L, Vaičiūnas G. Applying Analytic Hierarchy Process to assess traffic safety risk of railway infrastructure. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 376-383.
11. Büyüközkan G, Çifçi G, Güleryüz S. Strategic analysis of healthcare service quality using fuzzy AHP methodology. Expert systems with applications 2011; 38(8): 9407-9424, https://doi.org/10.1016/j.eswa.2011.01.103.
12. Büyüközkan G, Göçer F. An extention of ARAS methodology based on interval valued intuitionistic fuzzy group decision making for digital supply chain. 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Naples, IEEE, https://doi.org/10.1109/FUZZ-IEEE.2017.8015680.
13. Carpitella S, Certa A, Izquierdo J, La Fata C M. A combined multi-criteria approach to support FMECA analyses: A real-world case. Reliability Engineering & System Safety 2018; 169: 394-402, https://doi.org/10.1016/j.ress.2017.09.017.
14. Carpitella S, Mzougui I, Benítez J et al. A risk evaluation framework for the best maintenance strategy: the case of a marine salt manufacture firm. Reliability Engineering & System Safety 2020: 107265, https://doi.org/10.1016/j.ress.2020.107265.
15. Chua S J L, Ali A S, Alias A Bin. Implementation of Analytic Hierarchy Process (AHP) decision making framework for building maintenance procurement selection: Case study of Malaysian public universities. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(1): 7-18, https://doi.org/10.17531/ein.2015.1.2.
16. Ding S-H, Kamaruddin S. Maintenance policy optimization-literature review and directions. The International Journal of Advanced Manufacturing Technology 2015; 76(5-8): 1263-1283, https://doi.org/10.1007/s00170-014-6341-2.
17. Ding S-H, Kamaruddin S, Azid I A. Development of a model for optimal maintenance policy selection. European Journal of Industrial Engineering 2014; 8(1): 50-68, https://doi.org/10.1504/EJIE.2014.059349.
18. Doljak D, Stanojević G. Evaluation of natural conditions for site selection of ground-mounted photovoltaic power plants in Serbia. Energy 2017; 127: 291-300, https://doi.org/10.1016/j.energy.2017.03.140.
19. Emovon I, Norman R A, Murphy A J. Hybrid MCDM based methodology for selecting the optimum maintenance strategy for ship machinery systems. Journal of intelligent manufacturing 2018; 29(3): 519-531, https://doi.org/10.1007/s10845-015-1133-6.
20. Feng Z, Niu W, Wang W et al. A mixed integer linear programming model for unit commitment of thermal plants with peak shaving operation aspect in regional power grid lack of flexible hydropower energy. Energy 2019; 175: 618-629, https://doi.org/10.1016/j.energy.2019.03.117.
21. Fetanat A, Khorasaninejad E. Size optimization for hybrid photovoltaic-wind energy system using ant colony optimization for continuous domains based integer programming. Applied Soft Computing 2015; 31: 196-209, https://doi.org/10.1016/j.asoc.2015.02.047.
22. Fouladgar M M, Yazdani-Chamzini A, Lashgari A et al. Maintenance strategy selection using AHP and COPRAS under fuzzy environment. International journal of strategic property management 2012; 16(1): 85-104, https://doi.org/10.3846/1648715X.2012.666657.
23. George-Williams H, Patelli E. Maintenance strategy optimization for complex power systems susceptible to maintenance delays and operational dynamics. IEEE Transactions on Reliability 2017; 66(4): 1309-1330, https://doi.org/10.1109/TR.2017.2738447.
24. Ghosh D, Roy S. A decision-making framework for process plant maintenance. European Journal of Industrial Engineering 2009; 4(1): 78-98, https://doi.org/10.1504/EJIE.2010.029571.
25. Goossens A J M, Basten R J I. Exploring maintenance policy selection using the Analytic Hierarchy Process; an application for naval ships. Reliability Engineering & System Safety 2015; 142: 31-41, https://doi.org/10.1016/j.ress.2015.04.014.
26. Hashemi M, Asadi M, Zarezadeh S. Optimal maintenance policies for coherent systems with multi-type components. Reliability Engineering & System Safety 2020; 195: 106674, https://doi.org/10.1016/j.ress.2019.106674.
27. Hashemkhani Zolfani S, Bahrami M. Investment prioritizing in high tech industries based on SWARA-COPRAS approach. Technological and Economic Development of Economy 2014; 20(3): 534-553, https://doi.org/10.3846/20294913.2014.881435.
28. Ilangkumaran M, Kumanan S. Application of hybrid VIKOR model in selection of maintenance strategy. International Journal of Information Systems and Supply Chain Management (IJISSCM) 2012; 5(2): 59-81, https://doi.org/10.4018/jisscm.2012040104.
29. Jajimoggala S, Rao V V S K, Satyanarayana B. Maintenance strategy evaluation using ANP and goal programming. Management Theories and Strategic Practices for Decision Making, IGI Global: 2013: 117-138, https://doi.org/10.4018/978-1-4666-2473-3.ch007.
30. Jünger M, Liebling T M, Naddef D et al. 50 Years of integer programming 1958-2008: From the early years to the state-of-the-art. Springer Science & Business Media: 2009, https://doi.org/10.1007/978-3-540-68279-0.
31. Kato S, Fukumura N, Morito S et al. A Mixed Integer Linear Programming Approach to a Rolling Stock Rostering Problem with Splitting and Combining. RailNorrköping 2019. 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), Norrköping, Sweden, June 17th-20th, 2019, Linköping University Electronic Press: 2019; (069): 548-564.
32. Kim C S, Djamaludin I, Murthy D N P. Warranty and discrete preventive maintenance. Reliability Engineering & System Safety 2004; 84(3): 301-309, https://doi.org/10.1016/j.ress.2003.12.001.
33. Kiran Y V, Mahanty B. Reliability design of embedded systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2006; 30(2): 4-7.
34. Kurian M C, Shalij P.R, Pramod V.R. Maintenance strategy selection in a cement industry using analytic network process. Journal of Quality in Maintenance Engineering 2019; 26(4): 509-525, https://doi.org/10.1108/JQME-07-2017-0048.
35. Labib A W. A decision analysis model for maintenance policy selection using a CMMS. Journal of Quality in Maintenance Engineering 2004; 10(3): 191-202, https://doi.org/10.1108/13552510410553244.
36. Márquez A C. The maintenance management framework: models and methods for complex systems maintenance. Springer-Verlag London: 2007. doi:10.1007/978-1-84628-821-0, https://doi.org/10.1007/978-1-84628-821-0.
37. Martorell S, Villanueva J F, Carlos S et al. RAMS+ C informed decision-making with application to multi-objective optimization of technical specifications and maintenance using genetic algorithms. Reliability Engineering & System Safety 2005; 87(1): 65-75, https://doi.org/10.1016/j.ress.2004.04.009.
38. Nezami F G, Yildirim M B. A sustainability approach for selecting maintenance strategy. International Journal of Sustainable Engineering 2013; 6(4): 332-343, https://doi.org/10.1080/19397038.2013.765928.

39. Nguyen K-A, Do P, Grall A. Multi-level predictive maintenance for multi-component systems. Reliability engineering & system safety 2015; 144: 83-94, https://doi.org/10.1016/j.ress.2015.07.017.
40. Nguyen K-A, Do P, Grall A. Joint predictive maintenance and inventory strategy for multi-component systems using Birnbaum's structural importance. Reliability Engineering & System Safety 2017; 168: 249-261, https://doi.org/10.1016/j.ress.2017.05.034.
41. Özbek A. Performance Evaluation Of Turkey Diyanet Foundation By SAW, COPRAS and TOPSIS Method. Journal of Management and Economics Research 2017; 15(1): 66-84.
42. Özcan E, Danışan T, Eren T. A mathematical model proposal for maintenance strategies optimization of the most critical electrical equipment groups of hydroelectric power plants. Pamukkale University Journal of Engineering Sciences 2019; 25(4): 498-506, https://doi.org/10.5505/pajes.2018.38455.
43. Özcan E C, Danışan T, Yumuşak R, Eren T. An artificial neural network model supported with multi criteria decision making approaches for maintenance planning in hydroelectric power plants. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 21(3): 400-418, https://doi.org/10.17531/ein.2020.3.3.
44. Özcan E C, Ünlüsoy S, Eren T. A combined goal programming - AHP approach supported with TOPSIS for maintenance strategy selection in hydroelectric power plants. Renewable and Sustainable Energy Reviews 2017; 78(April): 1410-1423, https://doi.org/10.1016/j.rser.2017.04.039.
45. Panchal D, Chatterjee P, Shukla R K et al. Integrated Fuzzy AHP-CODAS Framework for Maintenance Decision in Urea Fertilizer Industry. Economic Computation & Economic Cybernetics Studies & Research 2017; 51(3) .
46. Panchal D, Kumar D. Maintenance decision-making for power generating unit in thermal power plant using combined fuzzy AHP-TOPSIS approach. International Journal of Operational Research 2017; 29(2): 248-272, https://doi.org/10.1504/IJOR.2017.083958.
47. Pourjavad E, Shirouyehzad H, Shahin A. Selecting maintenance strategy in mining industry by analytic network process and TOPSIS. International Journal of Industrial and Systems Engineering 2013; 15(2): 171-192, https://doi.org/10.1504/IJISE.2013.056095.
48. Rusin A, Wojaczek A. Improving the availability and lengthening the life of power unit elements through the use of risk-based maintenance planning. Energy 2019; 180: 28-35, https://doi.org/10.1016/j.energy.2019.05.079.
49. Saaty T L. The analytical hierarchy process, planning, priority. Resource Allocation. RWS Publications, USA 1980.
50. Sagar M K, Jayaswal P, Kushwah K. Exploring fuzzy SAW method for maintenance strategy selection problem of material handling equipment. International Journal of Current Engineering and Technology 2013; 3(2): 600-605.
51. Sankpal P, Andrew A, Kumanan S. Maintenance Strategies selection using Fuzzy FMEA and Integer Programming. Proceedings of the International Conference on Advances in Production and Industrial Engineering, 2015: 503.
52. Shafiee M. Maintenance strategy selection problem: an MCDM overview. Journal of Quality in Maintenance Engineering 2015; 21(4): 378-402, https://doi.org/10.1108/JQME-09-2013-0063.
53. Shafiee M, Labib A, Maiti J, Starr A. Maintenance strategy selection for multi-component systems using a combined analytic network process and cost-risk criticality model. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2019; 233(2): 89-104, https://doi.org/10.1177/1748006X17712071.
54. Shen L, Yang J, Jin X et al. Based on Delphi method and analytic hierarchy process to construct the evaluation index system of nursing simulation teaching quality. Nurse education today 2019; 79: 67-73, https://doi.org/10.1016/j.nedt.2018.09.021.
55. Shyjith K, Ilangkumaran M, Kumanan S. Multi-criteria decision-making approach to evaluate optimum maintenance strategy in textile industry. Journal of Quality in Maintenance Engineering 2008; 14(4): 375-386, https://doi.org/10.1108/13552510810909975.
56. Si T, Wang C, Liu R et al. Multi-criteria comprehensive energy efficiency assessment based on fuzzy-AHP method: A case study of post-treatment technologies for coal-fired units. Energy 2020: 117533, https://doi.org/10.1016/j.energy.2020.117533.
57. Taha H A. Integer programming: theory, applications, and computations. Academic Press: 2014.
58. Vahdani B, Hadipour H. Extension of the ELECTRE method based on interval-valued fuzzy sets. Soft Computing 2011; 15(3): 569-579, https://doi.org/10.1007/s00500-010-0563-5.
59. Wang J J, Jing Y Y, Zhang C F, Zhao J H. Review on multi-criteria decision analysis aid in sustainable energy decision-making. Renewable and Sustainable Energy Reviews 2009; 13(9): 2263-2278, https://doi.org/10.1016/j.rser.2009.06.021.
60. Wang Y, Deng C, Wu J et al. A corrective maintenance scheme for engineering equipment. Engineering Failure Analysis 2014; 36: 269-283, https://doi.org/10.1016/j.engfailanal.2013.10.006.
61. Wang Y, Liu Z, Liu Y. Optimal preventive maintenance strategy for repairable items under two-dimensional warranty. Reliability Engineering & System Safety 2015; 142: 326-333, https://doi.org/10.1016/j.ress.2015.06.003.
62. Zaim S, Turkyılmaz A, Acar M F et al. Maintenance strategy selection using AHP and ANP algorithms: a case study. Journal of Quality in Maintenance Engineering 2012; 18(1): 16-29, https://doi.org/10.1108/13552511211226166.
63. Zavadskas E K, Kaklauskas A, Sarka V. The new method of multicriteria complex proportional assessment of projects. Technological and economic development of economy 1994; 1(3): 131-139.
64. Zhang C, Gao W, Guo S et al. Opportunistic maintenance for wind turbines considering imperfect , reliability-based maintenance. Renewable Energy 2017; 103: 606-612, https://doi.org/10.1016/j.renene.2016.10.072.



Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Krešák J, Peterka P, Ambriško L, Mantič M. Friction lining coefficient of the drive friction pulley. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 338–345, http://doi.org/10.17531/ein.2021.2.13.

Friction lining coefficient of the drive friction pulley

Indexed by:

Jozef Krešák[a], Pavel Peterka[a], Ľubomír Ambriško[a], Martin Mantič[b]

[a] Faculty of Mining, Ecology, Process Control and Geotechnology, Technical University of Kosice, Park Komenskeho 14, 042 00 Kosice, Slovak Republic
[b] Faculty of Mechanical Engineering, Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic

Highlights

- The compare tests of the friction linings of mining hoisting system.
- The friction coefficient stability tests of the friction linings.
- The friction coefficient dependences on the weather conditions were analysed.
- Hardness was chosen as a representative property of the tested friction lining.

Abstract

Mine hoisting KOEPPE system or friction hoist winch work with traction pulley, the pulley rim grooves are lined. Lining has to provide a higher friction coefficient between the rope and the traction pulley. The constructors of mine hoisting machines require from the manufacturers a guaranteed appropriate and stabile value of a friction coefficient at different pressures between a rope and a friction lining under different external conditions (drought, moisture, icing, etc.). The paper presents processed measurements performed on the six samples of the friction lining (G1-G6) made of rubber and the sample of the standard used friction lining (K25). The samples (G1-G6) differ in the chemical composition of the rubber. Due to the confidentiality of the material composition of the friction linings the hardness of the lining material as a discriminator was chosen. The measured values of the friction coefficient of the rubber friction lining samples were compared with the values of the friction coefficient of the friction lining (K25) usually mounted on friction lining pulley.

Keywords

hoisting rope, friction pulley, friction lining, sliding friction coefficient.

1. Introduction

A mine hoisting system connects underground excavating areas with surface technologies; it transports extracted material, mine workers, machinery and equipment for exploitation. The main working element of mine hoisting machines is a steel wire rope, which can be wound on a drum or it can passed through a friction traction pulley. During any operation a drum hoisting system wind up one end of a hoist rope on the drum and a transport container via a cage suspension gear is gripped at the other end of the hoist rope. The rope is usually wound on the drum in one layer, but in the case of deep shafts the rope can be wound on the drum in two layers. The drum mining machines are double-acting, then the drum is divided and two hoist ropes are wound on it in two layers - upper and lower. The second frequently used type of a mine hoisting equipment is a machine with a friction pulley. These machines work with a rope passing through a friction pulley KOEPPE system or friction hoist winch, where the transport vessel travels between two horizons [5, 7]. Mańka et al. specified work of mining shaft hoist, depending on the drive type: in drum drives (rope is working in the underlap or overlap arrangement) or in drives with the frictional contact (KOEPPE system) [18]. Shirong investigated the friction coefficients between the steel wire rope and Polyvinylchloride (PVC) lining [20] and the hoisting friction conditions in a mine. The measurement shown: the friction coefficient decreases with increasing velocity or pressure and distribution of friction coefficients have a log-normal distribution [20]. Chang et al. studied wear and friction characteristics of the steel wire rope and the evolution of the tribological parameters at different friction stages [6]. Guo et al. based that force direction is deflected radially to the right. Force can be distributed into normal and friction force [11]. Ma and Lubrecht studied the local contact pressure between friction lining and steel wire rope. They developed first a 2-dimensional multi-grid code based on the geometry of steel wire rope [17]. Guo et al. investigated connection between friction transmission and longitudinal rope dynamics [12]. Zhang observed when steel wire rope is working around nylon pulleys; the bending fatigue life of steel wire ropes is twice longer than that of ropes working around steel pulleys [25]. In this article we describe and compare the linings of pulleys made of rubber and plastic. Standardized method [21] described in woks [1, 10] were used in the lining hardness tests. The utilization of the new lining material and development of the new lining construction lead to optimal repair maintenance [14], higher operation reliability and long life operation of the lining [15]. Material used for manufacturing of the friction lining requires high wear resistance [8] and on the other hand high friction coefficient on the contact with steel ropes. Rubber and plastic materials used for the manufacturing of the friction linings bring specific material properties [2] proper for specific operation condition of the mining hoisting system and especially for

E-mail addresses: J. Krešák - jozef.kresak@tuke.sk, P. Peterka - pavel.peterka@tuke.sk, L. Ambriško - lubomir.ambrisko@tuke.sk, M. Mantič - martin.mantic@tuke.sk

the friction lining pulley. The most important property of the friction lining is the stability of the friction coefficient under different weather conditions and the pressure between the rope and the groove of the pulley. Reliability and safety operation of the mining hoisting system depends on the optimization of above mentioned material properties and good friction properties of the chosen material [16].

2. Theory and calculation

The transmission of force from the drive friction traction pulley to the hoist rope is performed by means of friction between the rope and the groove of the rope pulley.

The drive wheel groove is often lined with a material having a higher coefficient of friction than steel. This material called lining is inserted into the rim of the drive pulley (Fig. 1). The magnitudes of tensile forces and a circumferential force on a friction disc or a drum are described by the Euler's equation.

When F + dF > F (Fig. 1b) than the resultant horizontal force dF_H is generated and if the rope is not to slide on the pulley than this force must be in equilibrium with the friction force dT, i.e. $dF_H = dT = dN.f$. Under such a condition the circumferential driving force is transmitted from the pulley to the rope without slippage according to the principle of belt friction. The force equilibrium in the horizontal direction of the x-axis is applied due to the distribution of forces in a rectangular coordinate system:

$$dF_H = (F + dF).\cos\frac{d\alpha}{2} - F.\cos\frac{d\alpha}{2}\ , \tag{1}$$

after the adjustment:

$$dF_H = F.\cos\frac{d\alpha}{2} + dF.\cos\frac{d\alpha}{2} - F.\cos\frac{d\alpha}{2}\ . \tag{2}$$

It is possible to consider: $\cos\frac{d\alpha}{2} \approx 1$ for a differentially small angle then:

$$dF_H = dF \tag{3}$$

The force equilibrium in the vertical direction of the y-axis (Fig. 1b) is given by the equality:

$$dN = (F + dF).\sin\frac{d\alpha}{2} + F.\sin\frac{d\alpha}{2}\ , \tag{4}$$

after adjustment:

$$dN = F.\sin\frac{d\alpha}{2} + dF.\sin\frac{d\alpha}{2} + F.\sin\frac{d\alpha}{2} \tag{5}$$

Concerning the differentially small angle it can be speculated that:

$$\sin\frac{d\alpha}{2} \approx \frac{d\alpha}{2}\text{, and}\quad dF.\sin\frac{d\alpha}{2} = 0\quad\text{then:}$$

$$dN = 2.F.\frac{d\alpha}{2} = F.d\alpha\ . \tag{6}$$

Considering that: $dF_H = dT$ and $dT = dN . f$; with both $dF_H = dF$ and $dN = F.\, d\alpha$, than after the substitution:

$$dF = f.F.d\alpha \tag{7}$$

This differential equation is solved by the separation of the variables and subsequently by their integration within the defined limits of the integration variables. To be aware of the fact that the force F increases along the pulley circumference from the smallest value F_2 to the largest value F_1, which corresponds with an increase of the wrap angle on the pulley from 0 up to the resulting angle α):

$$\int_{F_2}^{F_1}\frac{dF}{F} = \int_0^{\alpha} f.d\alpha\ , \tag{8}$$

then:

$$\ln\frac{F_1}{F_2} = f.\alpha\ , \tag{9}$$

and after delogarithmization a well-known Euler`s correlation arised (10):

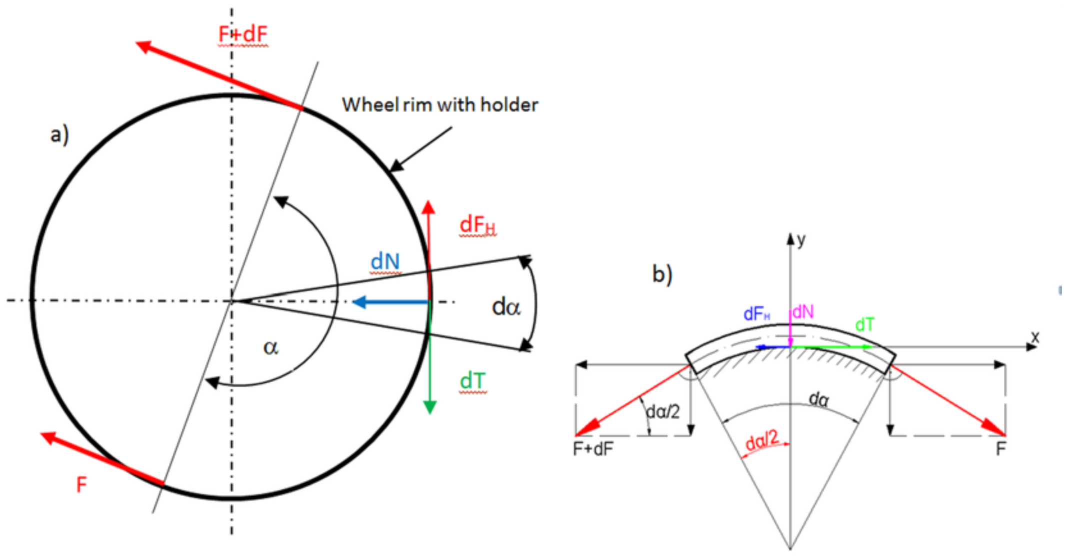


Fig. 1. Schematic force diagrams on the friction pulley a) schematic diagram of friction power transmission between friction lining and rope b) schematic diagram of friction power transmission between friction lining and rope on the elementary segment of the friction pulley

$$\frac{F_1}{F_2} = e^{f.\alpha} \tag{10}$$

During the actual hoisting the rope does not reach the state of gross slip (11):

$$F_1 = F_2.e^{f.\alpha} \tag{11}$$

where: F1 - tensile force in the digressional rope branch [N],
F2 - tensile force in the back-word running rope branch [N],
α - wrapping angle [rad],
f - friction coefficient [-].

3. Material and methods

3.1. Methodology for testing the coefficient of friction

According to the equation (11) there must be no slippage between the rope and the groove of the pulley friction lining, therefore it is necessary to know and guarantee the friction coefficient size between the rope and the friction lining. The methodology of friction force measurement is based on the principle of the equilibrium of tensile force and friction force (Fig. 1b). The tensile force (equal to the friction force) was recorded by a tensile tear tester. A measuring jig with lining samples together with a pressure force sensor was clamped in the tearing machine (Fig. 2a). According to this method the rope was inserted between two pieces of the same friction lining (Fig. 2a). The magnitude of the pressure force depends on the projection of the area of the pressed surface between the lining and the rope and on the required pressure (Fig. 2c).

The motion between the rope and the friction lining was caused by the pull of the testing machine. One of the jaws pulled the rope and the other one pulled the jig with the measured friction lining. To prevent the jig from sagging the jig arm is attached to the case by a ball joint. The case contains friction linings surrounding the rope from two sides. The rope axis passes through the imaginary axis of the jaws of the tearing machine and the box.

The force at which the motion between the rope and the friction lining took place was subtracted from the scale of the tearing machine. The starting force was considered to be the value valid for the coefficient of friction at rest (static coefficient of friction) and the force when moving was valid for the coefficient of friction in motion (dynamic coefficient of friction).

The test was performed on the non-lubricated rope and on the lubricated rope. The lubrication of the rope was performed in accordance with the Standard DIN 21258, i.e. the lubricant was applied to the rope and allowed to act at 20 °C for 16 hours. The test procedure was identical for both ropes. Due to the presence of water in mining environment the friction force between the lubricated rope and the friction lining groove was measured, while water was added to the contact area.

In the frame of the experiments the measurements were performed in order to determine the friction coefficient value for individual mixtures of the material used for the friction lining production. The friction lining samples used for the tests consisted of six types of rubber compounds (Fig. 3) and the K25 lining originally mounted on the traction pulley.

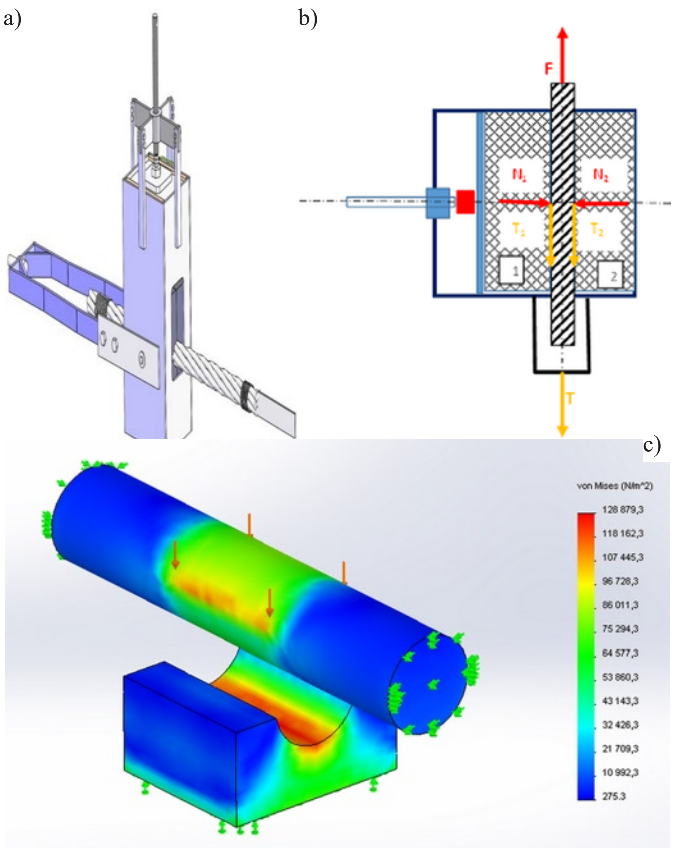


Fig. 2. The friction coefficient measurements, a) the measuring jig schema, b) the measuring jig schema with determination of N and T forces (right). c) the tension simulation model - friction lining and non-lubricated rope

The measurements were performed under the following conditions:

- friction lining and non-lubricated rope,
- friction lining and lubricated rope,
- friction lining, lubricated rope and water.

The values of the friction coefficient were measured at pressures: 1.5 MPa; 1.75 MPa; 2.0 MPa; 2.3 MPa between the rope and the friction linings.

For the calculation of the friction coefficient is valid (Fig. 2):

$$T = F \tag{13}$$

where: T – friction force [N],
F – tensile force recorded by the tear machine [N].

The equation for sliding friction is:

$$T = f.N \tag{14}$$

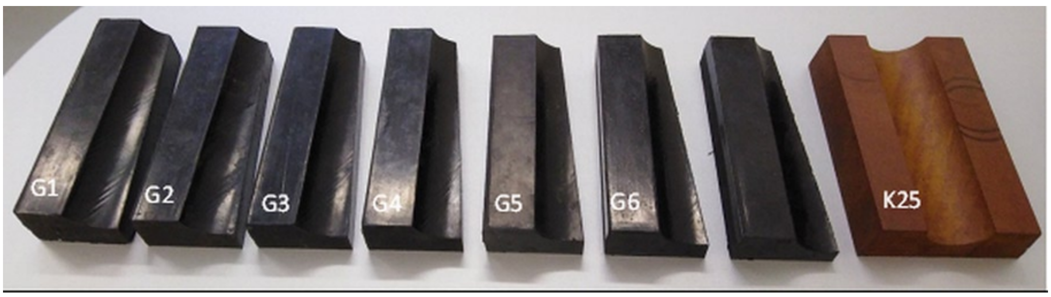


Fig. 3. Tested samples of the friction lining

where: N – normal force of pressure [N],
f – coefficient of sliding friction [-].

From the equilibrium of the forces according to the Fig. 2b follows:

$$T = T_1 + T_2 \tag{15}$$

where: T_1 – friction force from the lining No. 1 [N],
T_2 – friction force from the lining No. 2 [N].

If the equation (14) applies then:

$$T_1 = f.N_1 \tag{16}$$

$$T_2 = f.N_2 \tag{17}$$

where: N1 – force of the pressure on the friction lining No. 1 [N],
N2 – force of the pressure on the friction lining No. 2 [N].

The forces balance shown in the Fig. 2b:

$$N = N_1 = N_2 \tag{18}$$

After the substitution of the equations (16) and (17) in the equation (15) applies:

$$T = f.N_1 + f.N_2 \tag{19}$$

After the adjustment:

$$T = f.(N_1 + N_2) \tag{20}$$

If the equation (18) rates, then the equation (20) is:

$$T = f.(N + N) \tag{21}$$

The equation (22) for calculation of the friction coefficient is the result of the adjustment of the equation (21) and the use of the equation (13):

$$f = \frac{F}{2N} \tag{22}$$

3.2. The methodology of hardness testing

The tested rubber friction lining samples were divided following the hardness of the material. The hardness testing methodology is determined by the Standard STN EN ISO 868 and it specifies the method for determining the hardness of plastics and ebonite by indentation at which the depth of penetration of the tip is measured.

The Shore method is empirical; it is set for control purposes mostly. The hardness is inversely proportional to the tip intrusion. The tip intrusion depends on the modulus of elasticity and the visco - elastic properties of the tested material. The tip (made of hardened steel rod)

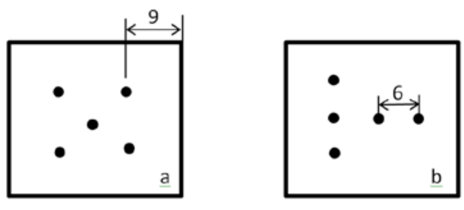


Fig. 4. Minimum distances in millimeters for the placemen of the punctures [1] A) distance of the punctures from the edge of the sample, B) distance between the punctures

has the shape of a beveled cone with an apex angle of 35° ± 0.25°, Ø 0.79 ± 0.03 mm [20].

The samples hardness was measured on five different places: the distance (Fig.4a) from the sample's edge min. 9 mm, the distance between punctures min. 6 mm (Fig. 4b) [1, 9].

4. Results

4.1. Measurement of the samples hardness

Hardness was chosen as a representative material property of the tested friction lining samples. The measured results (values for 5 punctures and average values of Shore hardness of rubber samples) are shown in the Table 1.

Table 1. Measured values of the Shore hardness of rubber samples G1 - G6 and sample K25

puncture/sample	K25	G1	G2	G3	G4	G5	G6
1	87,1	96,6	91,1	90,4	83	77	89,8
2	87,7	95,1	90,6	91,1	83	76,9	91,5
3	86,7	94,6	91,6	91,4	82,1	79,9	91,7
4	87,2	95,8	90,3	91,2	85,8	79,9	90,9
5	85,8	95	90,3	87,6	86,9	78,7	92,2
average	86,9	95,4	90,8	90,3	84,2	78,5	91,2

The maximum average value of the Shore hardness A / 15: 95.4 was measured out on the sample G1; the sample G5 showed the minimum average value of A / 15: 78.5.

The IRHD hardness measurements were also performed on the samples. The course of values of the hardness shows the Table 2. Both hardness measurements issue that the sample G5 has the softest material.

Table 2. Measured the IRHD hardness values of the rubber samples

sample	K25	G1	G2	G3	G4	G5	G6
IRHD	99,1	98,3	96,2	98,4	98,3	95,1	97,3

4.2. Measurement of the friction coefficient

The Fig. 5. shows the course of the friction coefficient measurement. The measuring jig was inserted between the jaws of the shredder. The rope sample was attached to one jaw and a lined fixture was attached to the other jaw. The defined pressure between the rope and the lining was developed by the pressure screw and then the moment, when it comes to the shift between the rope and the friction lining, was monitored.

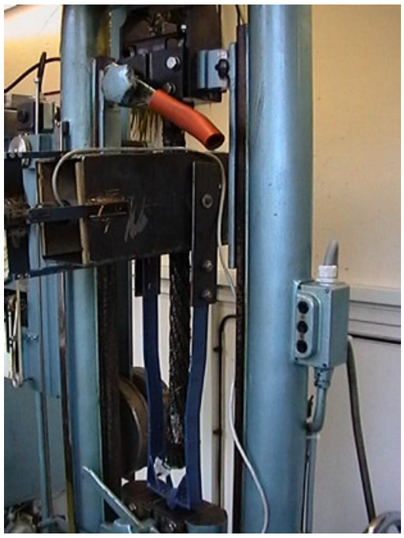


Fig. 5. The shredder and the measuring jig with the rope

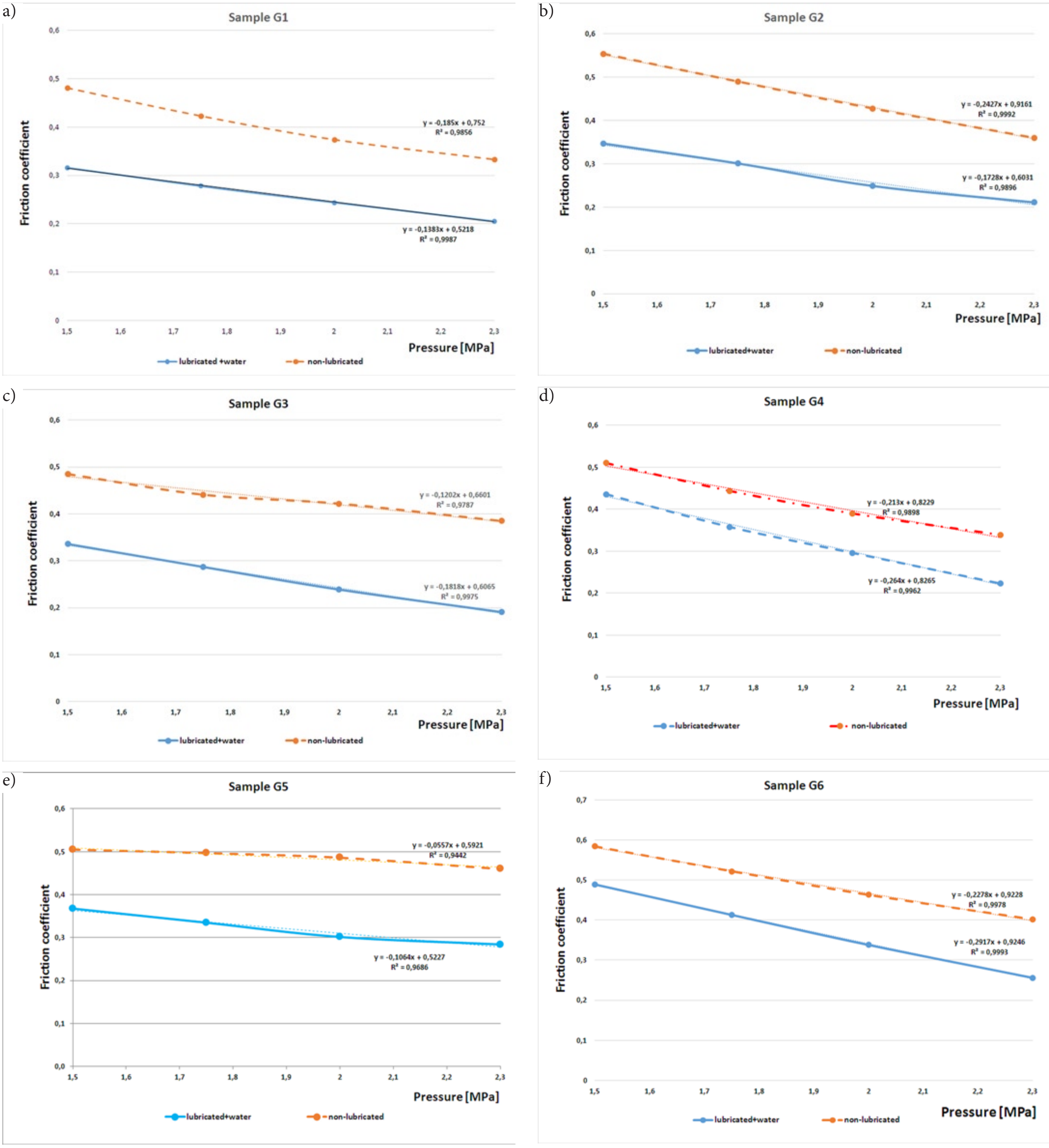
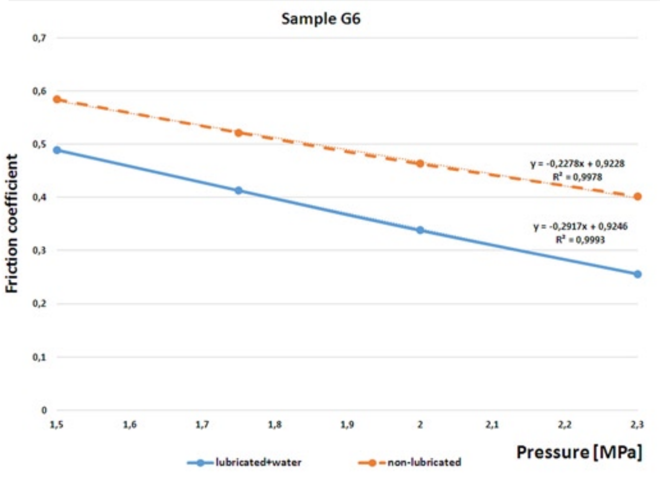


Fig. 6. The graphic dependences of the friction coefficient on the pressure magnitude between the lining and the rope; a– sample G1; b– sample G2; c– sample G3; d– sample G4; e– sample G5; f– sample G6; g– sample K25

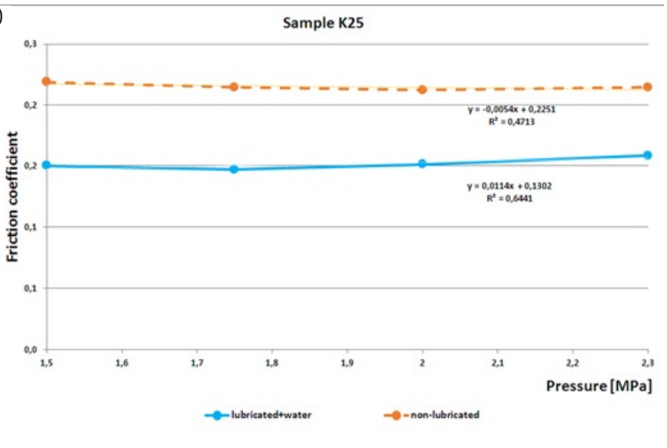
Fig. 6. a-i show the graphical dependences and the courses of the average values magnitude of the friction coefficients (from four measurements) depending on the magnitude of the pressure exerted between the rope and the friction lining, the condition of the ropes (non-lubricated, lubricated) and the presence of water.

5. Discussion

The rate of the perpendicular and parallel acting forces on the contact between two sliding bodies is defined as a friction coefficient [4]. Blau defined six categories of the testing devices for measuring of the friction coefficient in the laboratory condition [4]. Laboratory tests used for characterisation of the friction coefficient were made by many authors. They presented results of the basic friction tests made by gravitation based devices, direct linear force measurement devices or oscillation decrement devices. The test equipment used for above mentioned measurements can be defined as a tension wrap device according [4]. The measurement jig (Fig. 2) used for determination of the friction coefficient provide laboratory tests in the conditions close to real. Creep characteristics and dynamic friction transmission between friction lining and steel wire ropes were in situ investigated by Wang et al [23]. Wang et al. focused on correlation between effective g)



h)



i)

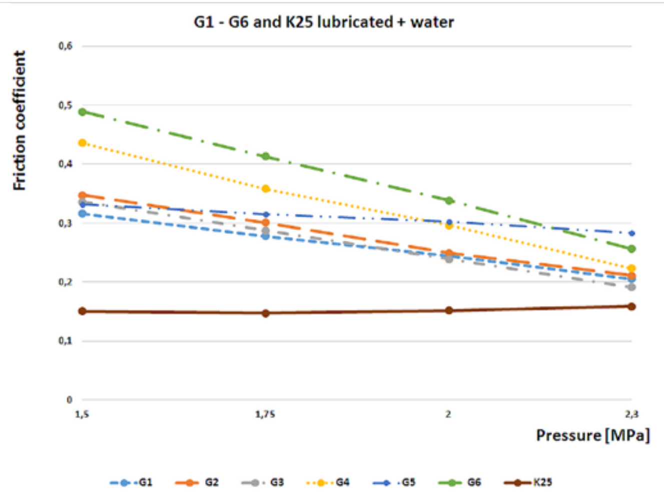


Fig. 6. [continued] g– sample K25; h – non lubricated: G1 - G6 and K 25; i– lubricated + water: G1 - G6 and K25

load, speed acceleration and deceleration to active slip angle, creep amplitude and creep velocity in process of vertical mining transport [23]. Dynamic contact characteristics between friction lining and steel wire rope were investigated by Wang et al. [22]. Wang et al. demonstrated effect of hoisting parameters on wear process of the lining and possibility of the gross slip [22]. Results presented in publications [22, 23] are focused on global effect of the hoisting process. Argatov and Chai studied effect of a variable friction coefficient on the fretting wear in conditions of gross sliding [3]. Argatov and Chai designed an asymptotic model for the progress of the contact area between the contacting surfaces [3]. The Figures 6 a-i are presented the measurements of the contact characteristics between the steel wire rope and the linings; friction coefficient of various rubber lining materials in various contact pressure is determined as well. Friction coefficients on the contact between the rubber block and smooth steel surface were investigated by Yamaguchi et al [24]. They focused on the effect of height and orientation of the rubber block on the friction coefficient [24]. Sliding friction characteristics of the water lubricated rubbers studied Ido et al. [13]. They focused on the surface topography of the rubber blocks and its effect on wet sliding friction characteristics [13]. Friction behaviour between glass plate and rubber was investigated by Nishi et al [19]. Rheological properties and effect of friction greases on friction between steel wire rope and fiction lining studied Feng at al. [9]. Feng et al. focused on the temperature and friction coefficient changes in the friction process and rheological properties of the friction-enhancing grease [9]. The results of the measurement manifest that regardless of the weather conditions (Fig. 6h and 6i) the friction lining samples made of rubber (G1-G6) have a higher friction coefficient than the sample K25. The samples comparison in terms of pressure between the rope and the friction lining indicates that the K25 friction coefficient does not change its value depending upon pressure (Fig. 6g). The course of the measured values of the friction coefficient of the friction lining K25 is much parallel to the pressure axis (Fig. 6g). The higher measured values of the friction coefficient for the friction linings G1 - G6 are evident pursuant to the comparison of all individual friction linings measurements (Fig. 6h and 6i). The significant changes in the values of the friction coefficient are visible, which in the case of friction linings G1 - G6 decreases considerably with increasing pressure (Fig. 6h, i). The decreasing course is identical with both ropes - non-lubricated and lubricated. From the samples G1 - G6 only the sample G5 has the course of the dependence of the friction coefficient on pressure similar to the sample K25, but the reached values of the friction coefficient are doubled.

The hardness values of the samples G4 and G5 are very similar to each other according to the Shore's hardness measurement method; the samples G4 and G5 are closest to the hardness value of the comparison sample K25 (Fig. 7).The sample G5 was chosen to be compared with the sample K25 due to the same trend of the course of the friction coefficient (Fig. 8, 9). Table 3 introduces directives and heading angles of the compared samples.

The trend lines overlayed by the values of the friction coefficient of the sample K25 with the non-lubricated and the lubricated rope are almost parallel to the x-axis (Fig. 8, 9). The sample made of K25 hardly changes the value of the friction coefficient depending on the value of the pressure between the friction lining and the rope. Likewise, the lines overlayed by the values of the friction coefficient of the sample G5 with the non-lubricated and the lubricated rope have the analogical course with the x-axis as well as the trend. The heading angles are

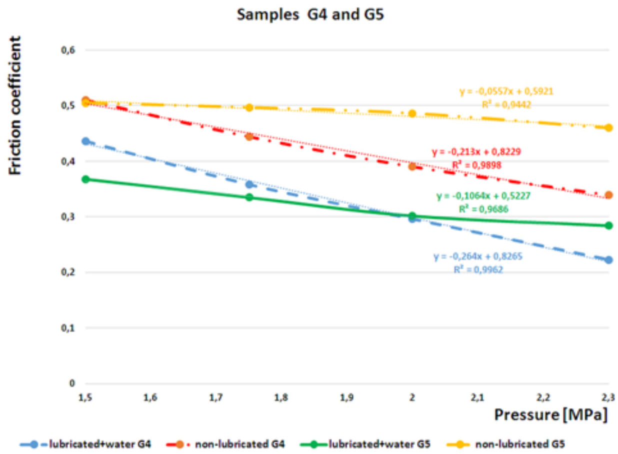


Fig. 7. Trends of the friction coefficient of the samples G4, G5 non-lubricated and lubricated + water

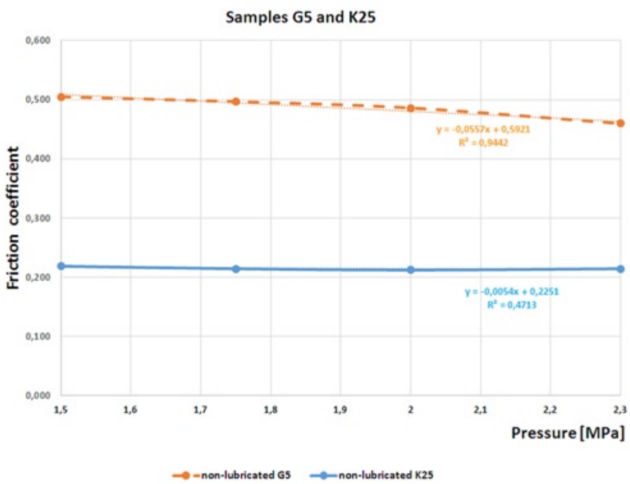


Fig. 8. Trends of the friction coefficient of the unlubricated samples G5 and K25

Table 3. Directives and heading angles α of the friction coefficient lines (Fig. 8 and 9)

Test condition	Sample K 25		Sample G5	
	Directive	Angle α	Directive	Angle α
non-lubricated rope	-0.0054	-0°18”	-0.0557	-3°11“
lubricated rope + water	0.0114	+0°39“	-0.1064	-6°4“

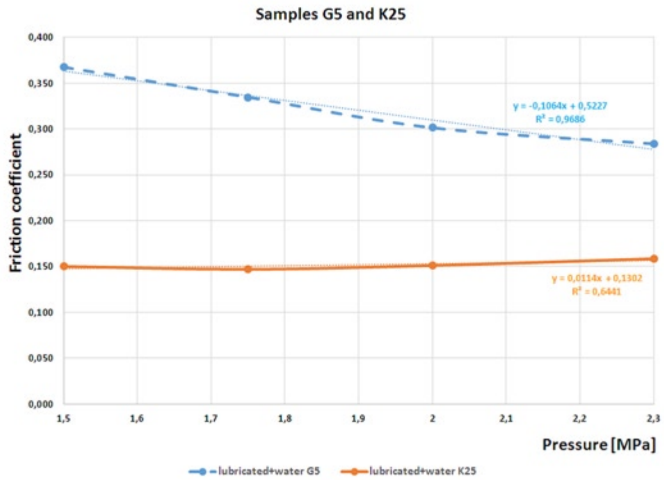


Fig. 9. Trends of the friction coefficient of the samples G5 and K25 lubricated +water

slightly larger than the angles of the sample K25, but they decrease with increasing pressure value (Fig. 8, 9).

6. Conclusion

The measured results of the friction coefficient show that all the rubber samples (G1 – G6) of the friction lining have a higher value of the friction coefficient than the friction lining K25. This applies to the entire pressure range between the rope and the friction lining, which is designated by the manufacturer of the towing equipment.

The trend of the values of the friction coefficient of the lining K25 is almost parallel to the axis of pressure (Fig. 6h and i, Fig. 8, 9). This means that the value of the friction coefficient changes very little with increasing pressure between the rope and the friction lining.

The friction lining G5 has a similar trend of the friction coefficient as the friction lining K25 (Fig. 8, 9). The rubber lining G5 shows a larger decrease of the friction coefficient value depending on the pressure than the lining K25. The values of the coefficient factor are significantly higher than the K25 has reached in the whole pressure range.

In terms of hardness the sample G4 is the closest one to the hardness values of the lining K25. Comparing it with the sample G5 (the lowest hardness of all rubber samples) the values of the friction coefficient of the sample G4 decrease significantly faster than it is with the sample G5 (Fig. 7).

Acknowledgments

Supported by the Scientific Grant Agency (VEGA) of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences, Grant No. VEGA 1/0075/20.

References

1. Ambriško Ľ. Determination of the abrasion resistance and the hardness of rubber covering layers. International Multidisciplinary Scientific GeoConference Surveying, Geology and Mining, Ecology and Management, SGEM conference proceedings, Bulgaria 2018; 18 (1.3): 255-262, https://doi.org/10.5593/sgem2018/1.3/S03.033
2. Andrejiova M, Grincova A, Marasova D. Analysis of tensile properties of worn fabric conveyor belts with renovated cover and with the different carcass type. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (3): 472-481, https://doi.org/10.17531/ein.2020.3.10.
3. Argatov I I, Young S. Chai Y S. Fretting wear with variable coefficient of friction in gross sliding conditions. Tribology International 2021; 153: 106555, https://doi.org/10.1016/j.triboint.2020.106555.
4. Blau P J. The significance and use of the friction coefficient. Tribology International 2001; 34: 585-591, https://doi.org/10.1016/S0301-679X(01)00050-0.
5. Chang X, Peng Y, Zhu Z, Gong X. Breaking failure analysis and finite element simulation of wear-out winding hoist wire rope. Engineering Failure Analysis 2018; 95: 1-17, https://doi.org/10.1016/j.engfailanal.2018.08.027.
6. Chang X D, Peng Y X, Zhu Z C, Zou S Y, Gong X S, Xu C M. Evolution Properties of Tribological Parameters for Steel Wire Rope under Sliding Contact Conditions. Metals (Basel) 2018; 8 (10): 743, https://doi.org/10.3390/met8100743.

7. Čereška A, Zavadskas E K, Bucinskas V, Podvezko V, Sutinys E. Analysis of Steel Wire Rope Diagnostic Data Applying Multi-Criteria Methods. Applied Sciences 2018; 8 (2): 260, https://doi.org/10.3390/app8020260.
8. Dzierwa A, Gałda L, Tupaj M, Dudek K. Investigation of wear resistance of selected materials after slide burnishing process. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (3): 432-439, https://doi.org/10.17531/ein.2020.3.5.
9. Feng C, Zhang D, Grecov D, Chen K. Effect of rheological properties of friction- enhancing greases on the friction between friction lining and wire rope. Tribology International 2020; 144: 106143, https://doi.org/10.1016/j.triboint.2019.106143.
10. Grinčová A, Marasová D. Experimental research and mathematical modelling as an effective tool of assessing failure of conveyor belts. Eksploatacja i Niezawodnosc - Mainteneance and Reliability 2014; 16 (2): 229-235.
11. Guo Y, Zhang D, Yang X, Feng C, Ge S. Experimental research on effect of wire rope transverse vibration on friction transmission stability in a friction hoisting system. Tribology International 2017; 115: 233-245, https://doi.org/10.1016/j.triboint.2017.05.033.
12. Guo Y, Zhang D, Chen K, Feng C, Ge S. Longitudinal dynamic characteristics of steel wire rope in a friction hoisting system and its coupling effect with friction transmission. Tribology International 2017; 119: 731-743, https://doi.org/10.1016/j.triboint.2017.12.014.
13. Ido T, Yamaguchi T, Shibata K, Matsuki K, Yumii K, Hokkirigawa K. Sliding friction characteristics of styrene butadiene rubbers with varied surface roughness under water lubrication. Tribology International 2019; 133: 230-235, https://doi.org/10.1016/j.triboint.2019.01.015.
14. Knopik L, Migawa K. Semi-Markov system model for minimal repair maintenance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (2): 256-260, https://doi.org/10.17531/ein.2019.2.9.
15. Kosobudzki M, Stańco M. Problems in assessing the durability of a selected vehicle component based on the accelerated proving ground test. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (4): 592-598, https://doi.org/10.17531/ein.2019.4.8.
16. Li Y H, Liang X J, Dong S H. Reliability optimization design method based on multi-level surrogate model. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (4): 638-650, https://doi.org/10.17531/ein.2020.4.7.
17. Ma W, Lubrecht A A. Detailed contact pressure between wire rope and friction lining. Tribology International 2016; 109: 238-245, https://doi.org/10.1016/j.triboint.2016.12.051.
18. Mańka E, Słomion M, Matuszewski M. Constructional Features of Ropes in Functional Units of Mining Shaft Hoist. Acta Mechanica et Automatica 2018; 12 (1): 66-71, https://doi.org/10.2478/ama-2018-0011.
19. Nishi T, Takeshi Yamaguchi T, Shibata K, Hokkirigawa K. Influence of unforced dewetting and enforced wetting on real contact formation and friction behavior between rubber hemisphere and glass plate during contacting and sliding processes. Tribology International 2020; 141: 105921, https://doi.org/10.1016/j.triboint.2019.105921.
20. Shirong G. The friction coefficients between the steel rope and polymer lining in frictional hoisting. Wear 1992; 152 (1): 21-29, https://doi.org/10.1016/0043-1648(92)90201-I.
21. STN EN ISO 868: 2003, Plastics and ebonite. Determination of indentation hardness by means of a durometer (Shore hardness).
22. Wang D, Li X, Wang X, Shi G, Mao X, Wang D. Effects of hoisting parameters on dynamic contact characteristics between the rope and friction lining in a deep coal mine. Tribology International 2016; 96: 31-34, https://doi.org/10.1016/j.triboint.2015.12.019.
23. Wang D, Zhang D, Mao X, Peng Y, Ge S. Dynamic friction transmission amd creep characteristics between hoisting rope and friction lining. Engineering Failure analysis 2015; 57: 499-510, https://doi.org/10.1016/j.engfailanal.2016.03.006.
24. Yamaguchi T, Katsurashima Y, Hokkirigawa K. Effect of rubber block height and orientation on the coefficients of friction against smooth steel surface lubricated with glycerol solution. Tribology International 2017; 110: 96-102, https://doi.org/10.1016/j.triboint.2017.02.015.
25. Zhang D, Chen K, Jia X, Wang D, Wang S, Luo Y, Ge Y. Bending fatigue behaviour of bearing ropes working around pulleys of different materials. Engineering Failure Analysis 2013; 33: 37-47, https://doi.org/10.1016/j.engfailanal.2013.04.018.

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Hryciów Z, Rybak P, Gieleta R. The influence of temperature on the damping characteristic of hydraulic shock absorbers. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 346–351, http://doi.org/10.17531/ein.2021.2.14.

The influence of temperature on the damping characteristic of hydraulic shock absorbers

Indexed by: Web of Science Group

Zdzisław Hryciów[a], Piotr Rybak[a], Roman Gieleta[a]

[a]Mechanical Engineering Faculty, Military University of Technology, Kaliskiego 2, 00-908 Warsaw, Poland

Highlights

- Shock absorber performance-temperature characteristic curves were determined.
- A method for determining the continuous use damper for ambient temperatures was proposed.
- Measurements over a wide temperature range showed profound changes in the damping factor.
- Energy dissipated during one cycle decreased linearly with the increase of the temperature.

Abstract

This paper presents the results of bench-tests and calculations assessing the influence of temperature on the performance of a two-pipe hydraulic shock absorber. The shock absorber prepared for the tests was cooled with dry ice to a temperature corresponding to that associated with the average winter conditions in a temperate climate. The temperature range of the shock absorber during testing was ensured via equipping it with a thermocouple and monitoring it with a thermal imaging camera. During testing, the shock absorber was subjected to kinematic forces of a selected frequency with two different, fixed displacement amplitudes. The results of the tests showed a direct correlation between the decrease of component resistance at lower temperatures. The rate of change in resistance was higher at lower temperatures. It was also found that the energy dissipated in one shock cycle decreased linearly with an increasing temperature. Finally, a method for determining the ideal use temperature of the shock absorber for the assumed operating conditions was also presented.

Keywords

hydraulic shock absorber, performance, damping factor, temperature, dissipation energy.

1. Introduction

Nowadays, in order to obtain the required damping values and the desired vibration response of car bodies, and the controlled response on its unsprung mass, double-acting hydraulic shock absorbers are usually implemented in unison with suspension spring elements. These shock absorbers are characterized by their different, two-way resistive forces during their compression stroke and their expansion stroke (the approach and distancing of the wheel to the body, respectively). The magnitude of these resistance forces depends on the parameters of the throttling valves used to decrease the flow of liquid between the chambers in the damper. The shock absorber functions via the conversion of the mechanical energy, generated by the transfer of liquid through the throttling valves, into the thermal energy.

Due to being a fundamental component for the comfort and safety of vehicle users, hydraulic shock absorbers are expected to meet high design, technical and operational requirements. These can be summarized as follows:

- stability of the damping characteristic curves over the assumed service life (or mileage),
- high operational efficiency under expected use conditions,
- as long a shock length as possible,
- low sensitivity to environmental factors (e.g., temperature, humidity),
- high mechanical strength and shock resistance,
- high durability, low weight and compact dimensions.

The correct functioning of shock absorbers is of particular importance for truck and off-road special-purpose vehicles. This is due to these vehicles travelling on poor quality or unpaved roads (where pits, thresholds, bumps and other irregularities appear), in various climatic and meteorological conditions, for a significant part of their service life. Under these conditions, the shock absorbers are then subjected to intense and complex loads with large displacement amplitudes and velocities. This intensive service life accelerates the wear of the shock absorbers components, and in turn, decreases their performance. As a result, the deterioration of the chassis and the vehicle steering are greatly expedited, in turn, affecting both the user comfort and safety.

Malfunctions of the shock absorber can be caused by, among other things, the following reasons: failure of the seals; too low level of the working fluid; valve leaks and vibrations; wear of the working elements; improper fastening (due to the loosening of the fixing elements); wear or loss of the rubber sleeve properties. From the development of these impairments, it is then possible to notice a progressive increase in the operating temperature of the shock absorber, significantly exceeding any recommended operating temperatures.

These permissible operating temperature ranges are specified by the shock absorber manufacturers. In most cases, they are in the range of −40 to 130°C [3], although for special applications (e.g., off-road vehicles) the maximum permissible temperature can reach up to

E-mail addresses: Z. Hryciów - zdzislaw.hryciow@wat.edu.pl, P. Rybak - piotr.rybak@wat.edu.pl, R. Gieleta - roman.gieleta@wat.edu.pl

180°C. This value is related to, among other characteristics, the type of seals used and their resistance to high temperatures. Temperature also affects the viscosity of the oil, and thus affects the damping characteristic curve of the shock absorber (i.e. the force-temperature relationship of the component). Reducing the viscosity of the oil when the temperature of the shock absorbers increases, as well as the pressure inside its chambers, in extreme cases can lead to the oil leakage. As a result, this causes the deterioration of the shock absorber's operation, or even irreversible damage. In conventional shock absorber designs, mineral oils are used as the working fluid. Whereas, in heavy-duty shock absorbers, synthetic oils are used, which are characterized by their greater resistance and decreased temperature-related viscosity changes. The viscosity of the oil used must not be too high, since this leads to problems at low operating temperatures (generating excessive force values that can lead to damage of the shock absorber). On the other hand, too low a viscosity increases the foaming tendency of the oil, reducing the damping forces and the lubricating properties. The typical dynamic viscosity of the mineral oils is around 40 MPa·s at 15°C [3].

In order to better identify the response of shock absorbers over time, a number of scientific publications have been focused on bench-tests and numerical calculations for the shock absorbers. These studies referred to the determination of the damping characteristic curves, the effect of the damping values on the vehicle dynamics [4, 19], and issues related to their mathematical modelling. In [16], for example, the authors presented a method for modelling a shock absorber (in particular, for its valves), and then confirmed their results based on the experimental results.

In previous scientific research, shock absorber tests results could be found that take into account temperature [5, 10, 14, 18] and their impact on the driving comfort of the user [13]. Studies of the shock absorbers filled with numerous types of fluids were also presented in [12], where the authors presented their heating characteristics during operation with a constant amplitude and frequency. This type of test determined the amount of heat exchanged between the shock absorber and the environment, and can be considered important for preventing the shock absorbers from overheating and the unfavorable changing of the damping characteristic curves. In [6] an attempt was made to assess the technical condition of vehicle-mounted shock absorbers based on their temperature changes during the vehicle movement. The temperature effect on the damping performance of the shock absorber, by implementing the Eusam method, was presented in [8]. The tests were carried out for two temperatures (−5°C and 20°C) and five different types of cars. The tests performed showed a decrease in the effectiveness of the shock absorbers in the range 5 to 25% at warmer temperatures. Changes in performance due to temperature changes were also analyzed for aircraft shock absorbers. Where, in [15], the authors presented the results of their experimental studies conducted at temperatures of −25°C, 0°C, 25°C and 50°C. Measurements of the shock absorber's damping characteristic curves were performed, as well as viscosity tests for hydraulic aviation oil. In this study, for positive temperatures (Celsius), a close to linear characteristic curve was observed for the viscosity decrease with temperature. At negative temperatures, the viscosity increased rapidly. As a result, the damping factor of the shock absorber was directly affected. These tests were also supplemented by calculations based on the CFD method.

The influence of the temperature on the basic characteristic curves was also evaluated for shock absorbers with variable damping characteristic curves. In [7], the influence of the temperature on the characteristic curves of a magnetorheological fluid shock absorber was described. In addition to the working fluid itself, the temperature influence on the resistance of the coils used in the control track were also studied, which was found to affect the resistance values of the shock absorber. Similar considerations were presented in [11], where a temperature range of 25 - 70°C was studied. Furthermore, examples of analytical studies were found in [1, 2, 10]. In [2], the authors presented a thermodynamic model of the shock absorber. This was used to simulate the change in the temperature of the shock absorber during its operation (for different movement speeds), until the vehicle's continuous use temperature was reached. Thus, determining the thermal energy dissipation capabilities of the shock absorber, where these results were also compared with experimental results. In another study [1], a mathematical model of the shock absorber and the heat flow between its elements was explored. This approach made it possible to determine the values of the damping forces, as well as the thermal effects caused by the changes in the geometry of the shock absorber and the properties of its components. Based on these calculations, it was determined that the highest temperature during the operation depends on the working fluid and the elements in direct contact with it, in particular the seals of the component.

In most of the literature found, results were presented for the shock absorbers at temperatures above 20°C; due to typical vehicle operating conditions. However, no results could be found in the literature for lower testing temperatures, such as the working temperature range, as declared by the manufacturers. Therefore, the main aim of the present study was to determine the influence of the working temperature range on the hydraulic shock absorber and its damping characteristic curves and energy dissipation capabilities. In addition, the possibility of heat dissipation to the environment was also explored.

2. Methodology

For the determination of the influence of the working temperature range on the hydraulic shock absorbers, a hydraulic shock absorber designed for installation in the rear suspension of a Renault D280 truck (with a custom bodywork) was tested. For this study, a telescopic Monroe E532080 shock absorber was used for the suspension of the vehicle.

The vehicle's shock absorber tests were carried out on the Instron 8802 electromechanical apparatus, which enables both static and dynamic tests. During the tests, the forces applied were controlled via the displacement signal.

The test program, as explained in further detail below, included the determination of the force-displacement characteristic curves of the rear shock absorber (without a rubber bushing and without a metal-rubber bushing) for different force frequencies at the fixed shock absorber strokes and for the different shock absorber strokes at a fixed force frequency. Ten load loops were recorded for all tests to determine the basic characteristic curves of the shock absorber, measuring both stroke and force frequency. In the case of testing the influence of the temperature on the damping characteristic curves of the shock absorber, the force and displacement were measured continuously from the initial temperature until the temperature value stabilized (or until it reached the maximum permissible value of 100°C). During the tests, the temperature of the shock absorber's surface (in its central part) was recorded using a J-type thermocouple. Thermal images were also taken using a FLIR model 6000 camera every 30 seconds. A view of the stand with the mounted shock absorber was shown in Fig. 1.

3. Results

3.1. Basic characteristic curves of the shock absorber

In the first stage of the experimental testing, the damping characteristic curves of the shock absorber were determined at a constant ambient temperature of 26°C. During these tests, the amplitude of the displacement was changed from 20 to 100 mm, while maintaining a constant frequency of 0.1 Hz. For each variant, the force value and shock absorber's piston displacement were recorded accordingly. Fig. 2 summarizes the force-displacement and damping characteristics of the shock absorber. When the shock absorber was stretched, its resistance was around 10 to 15 times greater than during compression. The characteristic curves provided show the limit force values of the pressure valve (from around 4.7 kN). This corresponded to a relative

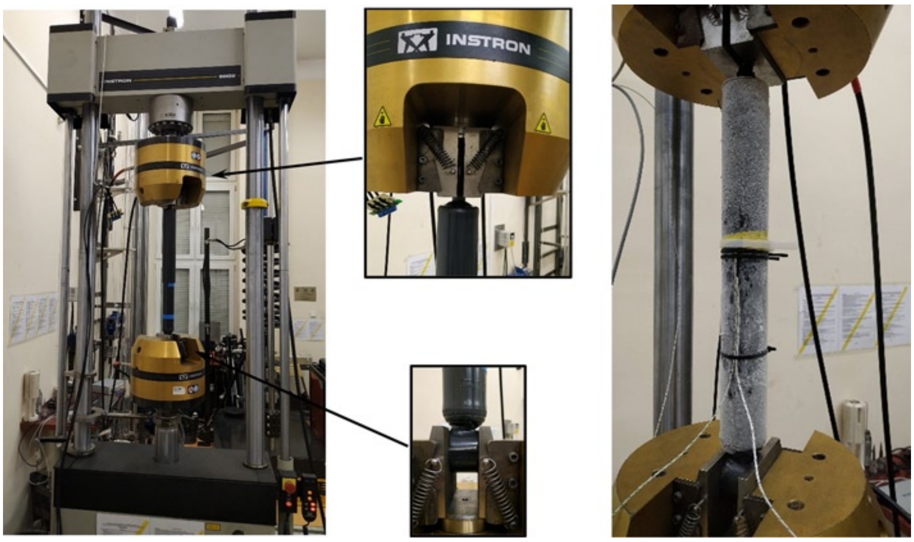


Fig. 1. The testing stand

piston speed of around 30 mm/s. For higher piston speeds, the force changed relative to the speed changes, and were smaller than in the initial force range. For the range of low speeds, the damping factor was 167 kNs/m, while at high speeds it was around 13 times lower reaching 12.9 kNs/m.

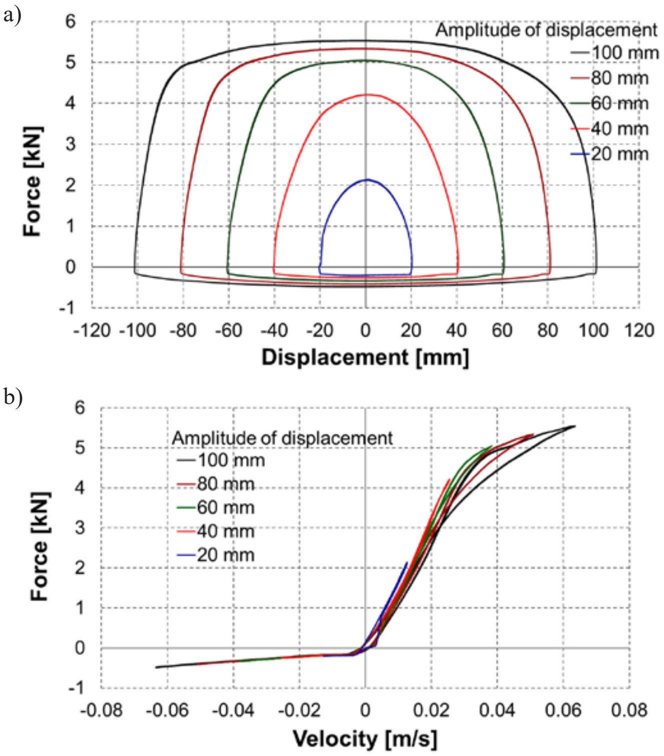


Fig. 2. Force-displacement and force-velocity characteristics of the shock absorber

3.2. Analysis of the heating process

An important issue for identifying the energy dissipation ability of the shock absorber was to study its heating and cooling processes. For this purpose, tests from the initial temperature to the operating temperature (state temps) were carried out for a fixed value of a displacement amplitude of 100 mm and a frequency of movement of 0.1 Hz. This temperature was measured on the outer surface of the shock absorber cylinder, thus being lower than the temperature of the oil inside the shock absorber. In addition, thermographic photographs were taken at fixed intervals (every 30 seconds). Fig. 3 shows the selected temperature distributions characterizing the process of heating the walls of the shock absorber. From this figure, it could be concluded that the cylinder heats up stronger in its upper part. The temperature difference between the upper and middle parts was around 5.1°C at the end of the measurements, and between the upper and lower parts as much as 16.3°C. Greater heating of the upper part was associated with, among other factors, the friction of the piston rod during its movement relative to the damper seal.

Fig. 4 displays the temperature changes found in the upper, middle and lower parts of the shock absorber (at points P1, P2 and P3 – Fig. 3, respectively). At the final stage of testing, the temperature on the surface of the shock absorber stabilized itself. Its increments were small and did not change by more than 0.1°C after an additional 60 s of the shock absorber's operation. Therefore, at this point, the study was terminated. The ability of the shock absorber to dissipate heat was characterized by, among others aspects, its geometry, the types of materials used or paint coatings applied. This process was described by the shock absorber cooling curve. In order to determine the shock absorber's cooling curve, the temperature changes of the heated shock absorber were measured after placing the subject in an environment with a constant temperature. The resulting curve was consistent with the examples found in the literature [14], and it should be emphasized that it depends on, among other things, the type of shock absorber and the prevailing environmental conditions (e.g., air temperature and humidity or relative air vlocity).

The change in temperature T(t) was described by the equation (1), expressing Newton's law of cooling [17]:

$$\frac{dT}{dt} = -k \cdot \left(T_p - T_{ot}\right) \tag{1}$$

where:

T_{ot} – ambient temperature [°C],
T_p – initial temperature [°C],
t – time [s],
k – decay constant [1/s].

Integration of equation (1) and rearrangement of terms leads to the expression in the form (2):

$$T(t) = T_{ot} + \left(T_p - T_{ot}\right) \cdot \exp(-k \cdot t) \tag{2}$$

Based on the approximation of the cooling results by function (2), the value of the decay constant k was determined to be around 0.034 s-1. The constant k can be described by the following relation (3):

$$k = \frac{\lambda \cdot S}{m \cdot c} \tag{3}$$

where:

λ – heat transfer coefficient [W/(m^2·K)],
S – heat exchange area [m^2],
m– mass [kg],
c – specific heat (J/(kg·K)).

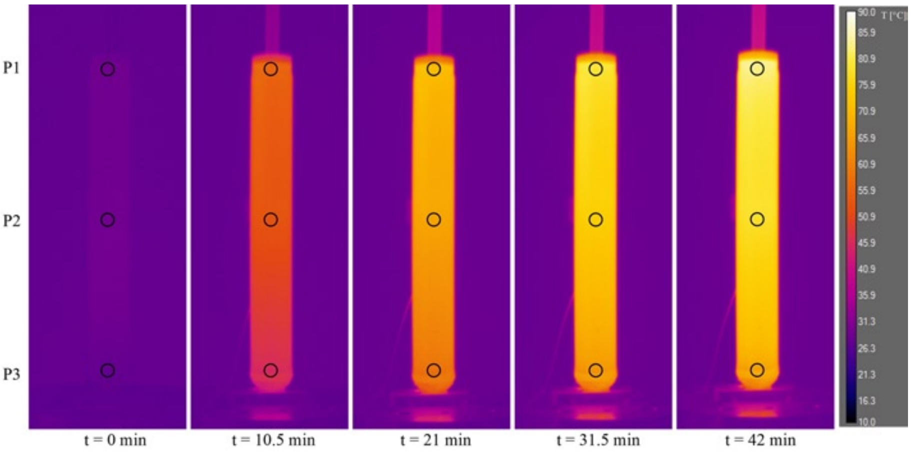


Fig. 3. Selected temperature distributions on the walls of the shock absorber

a)
b)

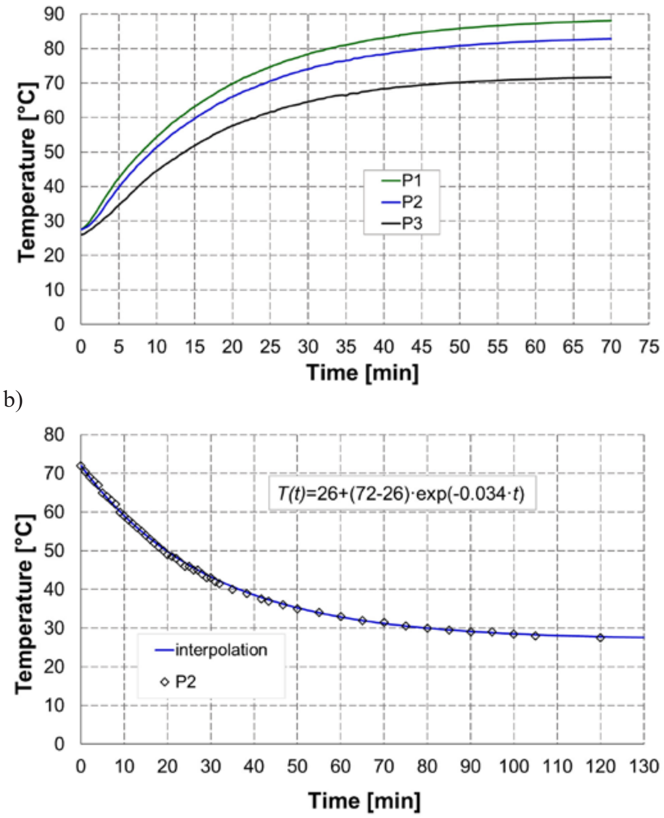


Fig. 4. Temperature changes on the surface of the shock absorber (left – heating, right – cooling)

Assuming the above-mentioned coefficients remain constant with respect to the temperature, it is possible (using Fourier's law for fixed heat flow conditions) to estimate the value of the continuous use temperature for other ambient temperatures. For this purpose, one should use the working conditions of the external force (energy dissipated in the shock absorber during one cycle) with the energy transferred to the external environment (4):

$$\dot{Q} = \frac{E_c}{t_c} = S \cdot \lambda \cdot \left(T - T_{ot}\right) \tag{4}$$

where:

$\dot{Q}$ – heat flux [W],
E_c – energy dissipated during one cycle [J],
t_c – duration of one operating cycle [s],

Transforming (4) yields relation (5):

$$S \cdot \lambda = \frac{E_c}{t_c \cdot \left(T - T_{ot}\right)} = const \tag{5}$$

For example, at an ambient temperature of 26°C, the tested shock absorber reached a continuous use temperature (in the middle part – point P2) of 83.71°C, dissipating an energy of 972 J during one cycle (lasting 10 s). If the amount of energy dissipated varies with temperature according to the dependence $E_c(T) = -2.15 \cdot T + 1152$ (fig. 7) at ambient temperature T_{ot2}, thus, the continuous use temperature T_2 is (6):

$$T_2 = \frac{S \cdot \lambda \cdot t_{c2} \cdot T_{ot2} + 1152}{S \cdot \lambda \cdot t_{c2} + 2.15} \tag{6}$$

At $T_{ot2} = -10$°C, after working with the same amplitude and frequency, the continuous use temperature of the shock absorber would reach $T_2 = 51.8$ ° C (dissipating an energy of 1040 J in one cycle). The presented method enables the evaluation of how appropriate the selected shock absorber for a specific car is for operation in the desired climatic conditions. According to literature, the increase in operating temperature in relation to the ambient temperature should not exceed 40-70°C [3].

3.3. Temperature influence on the damping characteristic curves

Given the final aim of the study was to determine the characteristic curves of the shock absorber at different operating temperatures. The tests previously described were carried out for two further displacement amplitudes of the shock absorber piston, while maintaining a constant frequency value of 0.1 Hz. The low value of the movement frequency permitted the minimization of the temperature difference between the oil and the walls of the shock absorber. In addition, the high speeds applied to the shock absorber lead to the generation of significant forces, exceeding the limit of opening the pressure limiting valve. Thus, causing a disturbance in the observation of the effect of temperature on the damping factor values. To reduce the heat exchange with the environment, the shock absorber cylinder was covered with a 20 mm thick layer of an insulating foam. This protected the cylinder walls against the intense heat exchange with the surroundings of the cooled shock absorber and against the heat transfer through the heated shock absorber. The insulation accelerated the heating process and enabled higher continuous use temperatures to be reached. For an amplitude of 50 mm, the tests were carried out in the temperature range from −40°C to 53°C (the established temperature value for the asserted conditions), and for 100 mm from −8°C to 100°C. Analyzing the results for the determination of the characteristic curves of the shock absorber at the different operating temperatures, it could be concluded that larger force differences in the shock absorber occur for smaller displacements (speeds), as seen in Fig. 5. For an amplitude of 100 mm the operation (i.e. opening) of the pressure limiting valve (and thus the force values) was observed for all temperatures. This was evidenced by the flattening of the upper part of the characteristic curves.

Fig. 6 provides a summary of the damping characteristics obtained in the temperature range from −40°C to 100°C. As previously mentioned, due to the rapid build-up of forces at −40°C and −20°C, the tests were carried out only for a displacement amplitude of 50 mm. Basing on these characteristics, the values of the damping factor were determined when the shock absorber was stretched in the low-speed range (before the pressure valve was opened).The results were summarized in Table 1. In addition, it shows the percentage changes in the

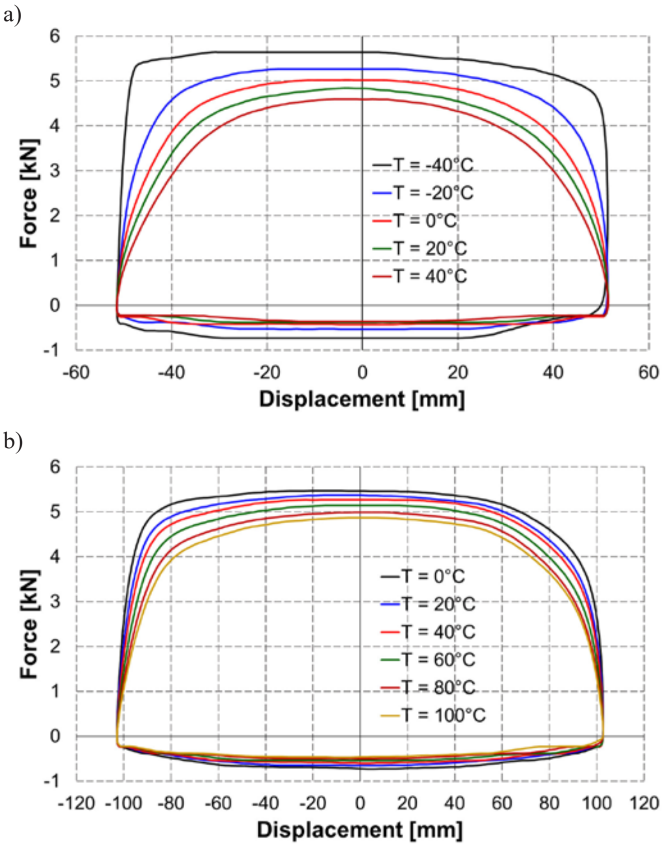


Fig. 5. Influence of the temperature on the force-displacement characteristics of the shock absorber (left – amplitude 50 mm, right – amplitude 100 mm)

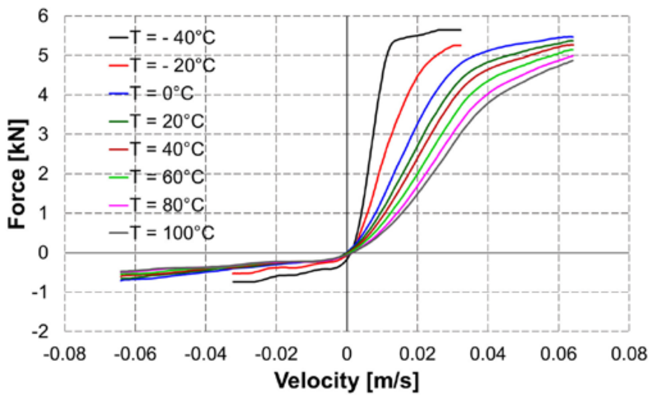


Fig. 6. Temperature influence on the damping characteristic curves of the shock absorber

Table 1. The changes of the damping coefficient in relation to a temperature

	Temperature [°C]							
	−40	−20	0	20	40	60	80	100
Damping coefficient c [kNs/m]	613.6	290.1	190.5	162.9	149.9	139.5	130.6	121.8
Relative difference $\delta c_{20°C}$ [%]	276.6%	78.1%	16.9%	0.0%	−8.0%	−14.4%	−19.8%	− 25.3%

value of the damping coefficient in relation to the value obtained at a temperature of 20°C.

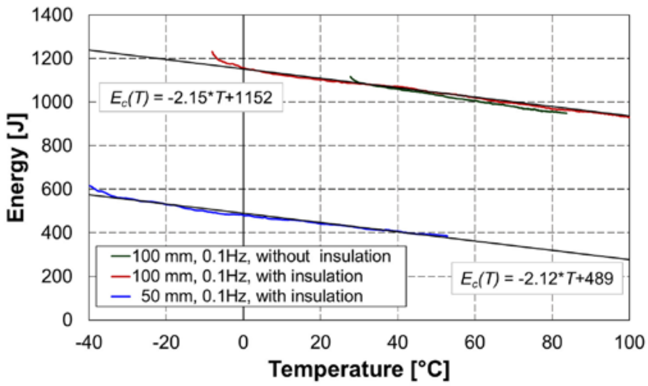


Fig. 7. Energy dissipated in one cycle

For the obtained changes in the damping force, as a function of temperature, the value of the dissipated energy was calculated for each full operating cycle as determined from the dependence (7):

$$E_c = \oint F dx \,, \tag{7}$$

For the initial part of the obtained characteristics curves, the curvature is well pronounced, as presented in Fig. 7. This was related to the time it took to transfer heat from the oil to the outer surface of the cylinder. After a short time, the heat transfer process was stopped. As a result, there was an almost linear decrease in the energy dissipated in one operating cycle with the increasing temperature.

For a force amplitude of 100 mm, the test was terminated when the temperature reached 100°C. For a force amplitude of 50 mm, only a temperature of 53°C was obtained. At this temperature, the amount of energy supplied to the system through the work of the external force evened the energy released to the environment. In addition, it could be stated that the rate of change in the amount of energy dissipated as a function of the temperature did not depend significantly on the displacement amplitude of the shock absorber piston. In both variants, a similar value was obtained (−2.15 J/°C for 100 mm and −2.12 J/°C for 50 mm). It also did not depend on the insulation used. For a shock absorber without insulation and covered with a layer of foam, a similar decrease in dissipated energy was observed with an increase in temperature.

4. Conclusions

This paper presents the results for laboratory bench-tests and calculations for determining the influence of temperature on the performance-based characteristic curves of a two-pipe hydraulic shock absorber installed on the rear suspension of a truck. Based on the results obtained, it was possible to conclude that the shock absorber's temperature had a significant influence on the damping factor values. This effect was particularly evident as the temperature decreased. Compared to the value obtained at 20°C, at 100°C the value of the damping coefficient decreased by around 25%, while at −40 ° C its value increased by around 280%.

Larger changes in the damping forces, due to temperature changes, were observed for smaller displacements of the shock absorber piston (lower movement speeds). For higher speeds, a pressure limiting valve was used.

Over a wide range of temperatures, the amount of energy dissipated during one cycle changed almost linearly with temperature. For the tested shock absorber, the specified rate of change was around −2.1 J/°C.

At low temperatures the relative motion of the elements of the shock absorber rapidly generated large resistance forces even at low speeds; with their value being limited by the pressure valve. High damping forces reduced the susceptibility of the suspension, increased the dynamic loads acting on the body and increased the likelihood of wheels tearing off the ground while driving.

References

1. Alonso M, Comas Á. Thermal model of a twin-tube cavitating shock absorber. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 2008; 222 (11): 1955-1964, https://doi.org/10.1243/09544070JAUTO829.
2. Cao J, Xie F, Ding E, Zhang X, Qian P, He K. Temperature Characteristics of the Valve-controlled Shock Absorber. IEEE 8th International Conference on Fluid Power and Mechatronics (FPM) 2019; 666-670, https://doi.org/10.1109/FPM45753.2019.9035857.
3. Dixon J C. The Shock Absorber Handbook, Second Edition, John Wiley & Sons 2007, ISBN: 978-0-470-51020-9.
4. Dukalski P, Będkowski B, Parczewski K, Wnęk H, Urbaś A, Augustynek K. Dynamics of the vehicle rear suspension system with electric motors mounted in wheels. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (1): 125-136, https://doi.org/10.17531/ein.2019.1.14.
5. Harkishandas R J, Gajjar S R, Patel A K. Experimental Analysis And Heat Transfer Study Of Damping Fluid In Shock Absorber Operation. International Journal of Engineering Development and Research (IJEDR) 2014; 2 (3): 2939-2947, http://www.ijedr.org/papers/IJEDR1403009.pdf.
6. Howard C, Sergiienko N Y, Gallasch G. Monitoring the age of vehicle shock absorbers. International Conference on Science and Innovation for Land Power 2018 (ICSILP 2018): 1-5.
7. Jastrzębski Ł, Sapiński B, Kozieł A. Automotive MR Shock Absorber Behaviour Considering Temperature Changes: Experimental Testing and Analysis. Acta Mechanica et Automatica 2020; 14 (1): 22-28, https://doi.org/10.2478/ama-2020-0004.
8. Jurecki R S. The influence of temperature on the damping value of shock absorbers determined by the Eusama method. Scientific Journals of the Maritime University of Szczecin 2019; 60 (132): 34-40.
9. Kubo P, Paiva C, Ferreira A, Larocca A. Influence of shock absorber condition on pavement fatigue using relative damage concept. Journal of Traffic and Transportation Engineering 2015; 2 (6): 406-413, https://doi.org/10.1016/j.jtte.2015.10.001.
10. Liang L, Liang T, Yunqing Z, Jie Z. Twin Tube Shock Absorber Thermo-Mechanical Coupling Simulation. Advanced Materials Research 2012; 566: 669-675, https://doi.org/10.4028/www.scientific.net/AMR.566.669.
11. McKee M, Gordaninejad F, Wang X. Effects of temperature on performance of compressible magnetorheological fluid suspension systems. Journal of Intelligent Material Systems and Structures 2018; 29(1): 41-51, https://doi.org/10.1177/1045389X17705203.
12. Patel D R, Rathod P P, Sorathiya A S. Heat Transfer Study of Damping Fluid and Improvement of Air-Gap in Shock Absorber Operation. International Journal Of Engineering Research & Technology (IJERT) 2012; 1 (3): 1-7.
13. Pavlov N. Influence of shock absorber temperature on vehicle ride comfort and road holding. MATEC Web of Conferences 2017; 133: 1-6, https://doi.org/10.1051/matecconf/201713302006.
14. Pracny V, Meywerk M, Lion A. Full vehicle simulation using thermomechanically coupled hybrid neural network shock absorber model. Vehicle System Dynamics 2008; 46 (3): 229-238, https://doi.org/10.1080/00423110701271864.
15. Shu N, Gu H, Liu H. Analysis of temperature effect on damping characteristics of landing gear shock absorber. International Conference on Aviation Safety and Information Technology (ICASIT 2020), 2020, https://doi.org/10.1145/3434581.3434593.
16. Skačkauskas P, Žuraulis V, Vadluga V, Nagurnas S. Development and verification of a shock absorber and its shim valve model based on the force method principles. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (1): 126-133, https://doi.org/10.17531/ein.2017.1.18.
17. Sobieski W. Termodynamika w eksperymentach, Uniwersytet Warmińsko-Mazurski 2015 Olsztyn.
18. Xie F, Cao J, Ding E, Wan K, Yu X, Ke J, Gao K. Temperature rise characteristics of the valve-controlled adjustable damping shock absorber. Mechanics & Industry 2020; 21 (111): 1-11, https://doi.org/10.1051/meca/2019084.
19. Yu B, Wang Z, Wang G, Zhao J, Zhou L, Zhao J. Investigation of the Suspension Design and Ride Comfort of an Electric Mini Off-Road Vehicle. Advances in Mechanical Engineering 2019; 11(1): 1-10, https://doi.org/10.1177/1687814018823351.
20. Yu Y, Zhao L, Zhou C, Yang L. Modelling and simulation of twin-tube hydraulic shock absorber thermodynamic characteristics and sensitivity analysis of its influencing factors. International Journal of Modeling, Simulation, and Scientific Computing 2018; 11 (2): 2050012-1 - 2050012-20, https://doi.org/10.1142/S1793962320500129.

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Kamiński M, Budzyński P, Hunicz J, Józwik J. Evaluation of changes in fuel delivery rate by electromagnetic injectors in a common rail system during simulated operation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 352–358, http://doi.org/10.17531/ein.2021.2.15.

Evaluation of changes in fuel delivery rate by electromagnetic injectors in a common rail system during simulated operation

Indexed by: Web of Science Group

Mariusz Kamiński [a],*, Piotr Budzyński [a], Jacek Hunicz [a], Jerzy Józwik [a]

[a] Faculty of Mechanical Engineering, Lublin University of Technology, ul. Nadbystrzycka 36, 20-618, Lublin, Poland

Highlights

- A high-intensity operation of electromagnetic injectors used in CR systems is simulated.
- Changes in fuel delivery rate by the injectors during the simulated operation are determined.
- Changes in the surface texture of fuel injectors control parts are evaluated.

Abstract

The objective of this study was to determine changes in fuel delivery rate by common rail system injectors during their simulated operation on a test stand. Four Bosch injectors used, among others, in Fiat 1.3 Multijet engines were tested. The injectors were operated on a test rig at room temperature for 500 hours (more than 72 million work cycles). During the test, pressure and injection frequency were changed. Changes in the operating parameters were estimated based on obtained injection characteristics and effective flow area determined thereby. The observed changes in fuel delivery rate were compared with results of the surface analysis of control valves and nozzle needles. Despite the stated lack of wear, significant changes in the dynamics of injector operation were observed, particularly at short injection times. Small pilot injections cannot be corrected by the fuel injection control system because they do not affect the changes in torque; however, they do affect the combustion process. This creates conditions for increased emission of toxic exhaust components.

Keywords

common rail, fuel delivery rate, injector operation.

1. Introduction

An internal combustion engine is currently the most widely used source of power for road transport. Given more and more demanding requirements for the emission of harmful exhaust components, engineers are striving to improve the combustion process in order to ensure the lowest possible emissivity and the highest possible thermal efficiency [28]. The fuel injection process and its control play a very important role in this respect [1, 5, 11].

Currently, compression ignition engines predominantly use common rail fuel injection. The design of such system makes it possible to minimize inaccuracies of fuel injection resulting from inertia of the injection system components, which is typical of systems based on injection pumps. Fuel pressure in the end part of the common rail system, i.e. the area between injectors and a high-pressure pump, is maintained at the same level, which allows for maintaining the same operating conditions for all engine cylinders. Since fuel is injected in several doses, it is possible to obtain the desired fuel distribution in the cylinder, which allows for the control of the combustion process and thus reduced emission of toxic exhaust gas components and lower pressure rise rate [22, 27]. However, this requires high precision of fuel metering and maintenance of constant fuel delivery rate characteristics during operation. The combustion process is particularly sensitive to fuel injection process and pilot injection quantity [9], therefore the metering of small fuel quantities in the ballistic regime for injector needles requires high precision.

Common rail fuel injection systems are often associated with various types of operational malfunctions. These predominantly result from damage of the injection system components, i.e. a high-pressure fuel pump or – more frequently – injectors. Injector damage is usually caused by poor quality of fuel [30, 31, 33] or cavitation erosion related to rapid flow of fuel through the injector [2, 13, 14, 32]. Poor fuel quality usually leads to accelerated accumulation of deposits; it can also accelerate erosive wear if fuel contains fractions with low vapour pressure.

Birgel et al. [3] reviewed the literature on the mechanisms of deposit formation and their effect on injector characteristics. Research has shown that fuel flow rate across the injector decreases linearly during operation, resulting in a corresponding reduction in engine power. Following 30 hours of engine operation at full load, the fuel delivery rate was reduced by about 3.5%. The literature review shows that many studies focus on the influence of biocomponents on deposit formation [15, 17]. In general, biofuels induce the formation of hard polymers inside the injectors and of carbon deposits around the injector holes. Hofmann et al. [7] investigated the effect of worn injectors on changes in control system signal characteristics, namely the needle lift and the fuel pressure in the feed line. It was found that a worn injector – its mileage unknown – had the needle lift reduced

(*) Corresponding author.
E-mail addresses: M. Kamiński - mariusz.kaminski@pollub.pl, P. Budzyński - p.budzynski@pollub.pl, J. Hunicz - j.hunicz@pollub.pl, J. Józwik - j.jozwik@pollub.pl

by about 20% compared to a new injector. A similar trend was observed with respect to the fuel delivery rate. The reduced needle lift and fuel flow rate were attributed to the formation of hard deposits. However, the authors observed that the reduced needle lift during operation is to some extent compensated for by cavitation wear of the injector holes.

Apart from deposit formation and cavitation erosion, the problem of tribological wear of injector components is often discussed in the literature too, especially in terms of alternative fuels. However, it should be emphasized that the available studies mainly concern older generation injection systems operating at pressures being an order of magnitude lower than those in common rail systems. Obtained results are not unequivocal, predominantly due to the limited time of the studies. Although visible changes in the surface profile are observed, they are not regarded as having direct impact on operating parameters of the fuel injection system components [25]. With good lubrication, geometrical changes in the interacting injection system components are negligible. The study conducted by Niewczas et al. [20] showed that after the durability test of 60 million cycles, the plunger diameter in the injection pump increased by about 4 μm.

Schuckert et al. [29] investigated the relationship between injector aging and fuel delivery rate. The results showed that reduced cross section of the injector holes due to deposit formation does not necessarily lead to reduced fuel delivery rate at short injection times. Hydraulic phenomena cause the injection time to be longer at the same energizing time, thus compensating for the lower fuel flow rate. Payri et al. [24] quantified the closing hydraulic delay of a piezoelectric injector. The aged injector was characterized by an increased fuel flow rate than the new injector for the same energizing time.

Optical analyses of fuel spray conducted by Hofmann et al. [7] showed that the reduced injector hole make the fuel spray narrower and increase its range, while the larger diameter holes have the opposite effect. Pielecha et al. [26] conducted a comparative analysis of fuel spray characteristics for new and aged injectors. Following a mileage of 80,000 km, the worn injectors were characterized by 30% reduced cone angles. A slight increase in the spray range was observed too. The above changes in fuel spray geometry indicate reduction in the cross section of the holes during operation.

Recently, Payri et al. [23] investigated the effects of injector aging on multiple injection strategies. The injector was subjected to wear on a test rig under controlled conditions. The researchers quantified the observed changes in the fuel flow rate in the injector, attributing them to reduced hole diameters. Similarly to Schuckert et al. [29], they found that increased injection timing is significant and that multiple injection leads to continuous fuel flow, as in the case with single injection.

Research has also shown that even new injectors differ significantly with respect to their flow characteristics. Ferrari et al. [4] determined the characteristics of fuel injection rate depending on the injection timing for different pressures in the fuel rail. Results obtained for the tested injectors showed that at short energizing times typical of pilot injection, the spread of injected fuel quantity is greater than the mean value. Changes in the injection characteristics and injection rate have a direct impact on the combustion process in the cylinder. The pilot injection quantity is of particular importance [8], and small fuel quantities are particularly exposed to variability resulting from wear and geometrical changes of the injectors.

Hofmann et al. [6] proposed an algorithm for injection strategy correction to compensate for changes resulting from injector aging. A simplified model of injection and combustion was sufficient to select optimal injection strategies. Importantly, the study found that a mere change of the injection time is not sufficient to compensate for changes in its characteristics, due to the fact that apart from the fuel flow coefficient, the spray characteristics and dynamic parameters are changed too. Nevertheless, in the study, only changes in the cross section of the holes were considered to be an indicator of wear.

The literature review shows that, on the one hand, detailed data are available about tribological processes and the relationship between technical condition and injection characteristics; on the other hand, however, the actual changes in fuel delivery rate characteristics are unknown. These gaps in knowledge result from the fact that the assessment of changes in the operating characteristics of injectors requires long-term studies under controlled conditions. In real operating conditions, simultaneously occurring processes, such as abrasive wear and internal deposits formation or cavitation wear of injector holes and carbon deposits formation, can compensate each other.

The objective of this study was to determine changes in injection characteristics of the analysed injectors following a high-intensity 500-hour laboratory durability test, amounting to a mileage of approx. 70,000 km. As a result, it was possible to determine the extent of changes in the injector characteristics during the run-in period. Tests were conducted at room temperature, which reduced, among others, the formation of internal deposits. In effect, it was possible to isolate the effects related only to tribological wear. The changes in injection rate characteristics were compared with changes in the surface texture of key elements of the injectors.

2. Simplified description of fuel injection process

In common rail systems, fuel quantity metering is adjusted by changing injector valve opening time. Theoretically, fuel flow through an open valve is steady and results from the fuel nozzle cross section and fluid velocity:

$$\dot{m}_{th} = \rho_f A_0 v_{th} \,, \tag{1}$$

where ρ_f is the fuel density, A_0 is the total cross-sectional area of the injector nozzle openings, and v_{th} is the theoretical flow velocity of fuel across the injector tip. The flow velocity is derived from the Bernoulli equation:

$$v_{th} = \sqrt{\frac{2\Delta p}{\rho_f}} \,, \tag{2}$$

where Δp is the difference between the pressure of injected fuel and the pressure across the area into which fuel is injected. However, the actual flow rate is corrected by the flow coefficient μ, defined as the ratio of actual to theoretical mass flow:

$$\mu = \frac{\dot{m}}{\dot{m}_{th}} \,. \tag{3}$$

Based on Equations (1) and (2) and taking into account the coefficient μ, the actual mass flow rate can be expressed as:

$$\dot{m} = \mu A_0 \sqrt{2\rho_f \Delta p} \,. \tag{4}$$

For convenience, the actual fuel spray cross section can be presented as the product of the geometric cross section of the valve and the flow coefficient. Additionally, by introducing the injector opening time t_r, based on the measured fuel volume V corresponding to one fuel injection, the actual cross section of the fuel spray can be calculated as:

$$\mu A_0 = \frac{V}{\sqrt{2\frac{\Delta p}{\rho}} t_r} \,. \tag{5}$$

However, the actual injector valve opening time in Equation (5) is not the same as the theoretical energizing time of the solenoid injector. The delay in injector opening results from a curve of inductor current increase, mechanical resistance, hydraulic resistance and inertia of both the fluid and the injector mechanical components. The delay occurs at injector closing too, due to the above reasons. It is worth noting that the delay values are so large that at short energizing times, the physical injection begins after the end of electric pulse [16]. Assuming that the delays are independent of the injected fuel quantity, the total difference Δt between the energizing time ET and the actual injection time t_r can be determined by analysing injection rate characteristic (Fig. 1). The intersection of the linear regression line with the abscissa indicates the time at which the theoretical injection rate is equal to zero. This time is the difference between the valve opening delay and the valve closing delay. Therefore, knowing the injection rate characteristics for a specific type and pressure of fuel and solenoid timing parameters, the actual injection time can be estimated based on the following dependence:

$$t_r = ET - \Delta t \ . \tag{6}$$

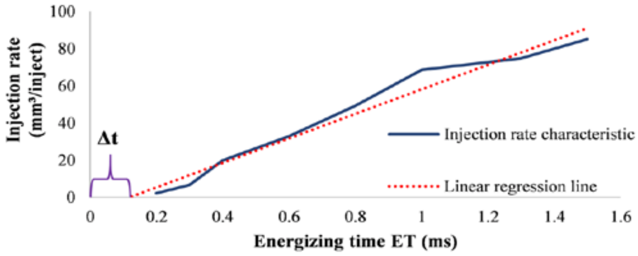


Fig. 1. Injector dead time determined based on injection characteristics

3. Research methods

Bosch high-pressure electromagnetic fuel injectors applied in common rail fuel injection systems were used in the study. Four injectors of the same type (0445110183) were tested. Laboratory tests were carried out on a test rig for testing fuel injection pumps and fuel injectors, STPiW-3, manufactured by Autoelektronika Kędzia (Fig. 2).

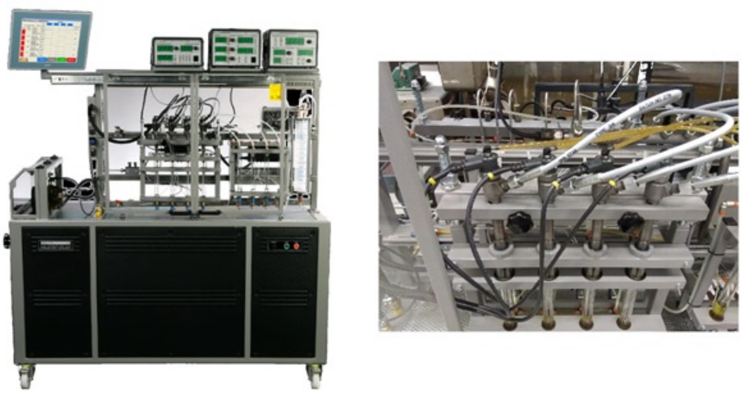


Fig. 2. STPiW-3 test rig for testing fuel injection pumps and fuel injectors

The injectors were operated for 500 hours, which amounted to about 72.18 million work cycles. To simulate operation of the injectors under conditions reflecting the real ones, the injection pressure and injection frequency were changed during the test. Detailed information about the experimental parameters is given in Table 1.

The fuel used in this study was Ekodiesel Ultra diesel oil produced by PKN ORLEN (Poland), the standard properties of which are given in Table 2. To ensure the same properties of the fuel, its temperature was stabilized at 40 ± 2°C.

Table 1. Parameters of performance tests

Work time (h)	Total work time (h)	Injection pressure p_{rail} (MPa)	Injection frequency (Hz)	Share in the test * (%)
120	120	100	25	15
30	150	120	35	5
150	300	120	40	30
130	430	120	50	33
70	500	140	50	17

* *share in the test denotes the percentage number of injections made under given operating conditions to the overall number of injections made by the injector during the test.*

Before and after the performance test, volumetric injection rate characteristics were determined. Conditions applied in the test are listed in Table 3. They were selected in such a way as to cover most operating conditions occurring in real conditions, i.e. low, medium and high injection pressures, as well as short and extended injection timing. The longest injection times of 1.5 and 2 ms were selected for the test due to the fact that they ensure a long period of stable mass flow rate, thanks to which it is possible to accurately determine the flow coefficient for the injector. All combinations of the variables were tested, yielding a total of 100 measurement points. The measurements were repeated five times, and based on obtained results, the average injection rates were calculated for every measurement point.

After the durability test, the injectors were tested on a specialized test bench EPS 945 from BOSCH, which is used for testing and issuing correction codes in accordance with the procedure specified by the manufacturer. Obtained changes in the injection rate were evaluated in terms of statistical significance. The evaluation was performed using Statistica 13.0. Since statistical features of the obtained results (distribution and variance in individual tests) precluded the use of parametric tests, the Wilcoxon signed-rank test was used instead. The level of statistical significance was set at $\alpha = 0.1$, which corresponds to the values used in technical sciences [12].

After that, the injectors were disassembled and topographically examined using the Alicona InfiniteFocus5G optical device for surface roughness and texture measurements. Surface roughness was measured in two stages. First, a 3D model was created using the InfiniteFocus technology. Next, a surface profile was extracted from the 3D model, and – on the basis of this profile – roughness parameters were calculated in compliance with ISO standards. In this way, it was possible to identify changes in the surface of key elements of the injectors, i.e. control valves and injector needles.

4. Results and discussion

In accordance with the assumptions of the experiment, changes in injector operation were evaluated predominantly based on the changes in their injection rate. Fig. 3 shows the characteristics of injection rate for selected parameters of their operation from the values given in Table 3.

An analysis of the above injection rate characteristics demonstrates that in almost every case, regardless of the injection pressure, the injection rate decreases during the test. A decrease in slope of the injection rate characteristics curve in relation to the horizontal axis of the graph, in accordance with the scheme shown in Fig. 1, indicates changes in the effective cross-sectional area of the injector, as illustrated by the characteristics shown in Fig. 4.

An analysis of the behaviour pattern of the effective cross-sectional area of the injector demonstrates that this parameter undergoes significant changes in the entire range of measurement points. Statistically significant changes can be observed for the average value of the effective cross-sectional area of the injector. According to the statistical test results, the value of p is 0.086 for Injector I, while for Injectors

Table 2. *Standard physical and chemical properties of the fuel used in the study, according to the data provided by the manufacturer*

Property	Unit	Specifications		Test
		Lower limit	Upper limit	
Cetane number	-	51.0	-	PN-EN ISO 5165
Cetane index	-	46.0	-	PN-EN ISO 4264
Density at 15 °C	kg/m^3	820.0	845.0	PN-EN ISO 12185 PN-EN ISO 3675
Flash point	°C	56.0	-	PN-EN ISO 2719
Carbon residue (on 10% distillation residue)	%(m/m)	-	0.30	PN-EN ISO 10370
Ash residue	%(*m/m*)	-	0.010	PN-EN ISO 6245
Water content	mg/kg	-	200	PN-EN ISO 12937
	% (m/m)	-	0.020	
Oxidation stability	g/m^3	-	25	PN-ISO 12205
	h	20.0	-	PN-EN 15751
Lubricity, corrected wear scar diameter (wsd 1.4) at 60 °C	mm	-	460	PN-EN ISO 12156-1
Viscosity at 40 °C	mm^2/s	2.000	4.000	PN-EN ISO 3104

Table 3. *Injector parameters used for determination of injection characteristics*

Parameter	Tested values
Injection pressure, p_{rail} (MPa)	30; 40; 50; 70; 80; 90; 100; 120; 150; 170
Nominal injection time, ET (ms)	0.2; 0.3, 0.4; 0.6; 0.8; 1.0; 1.3; 1.5; 1.8; 2.0

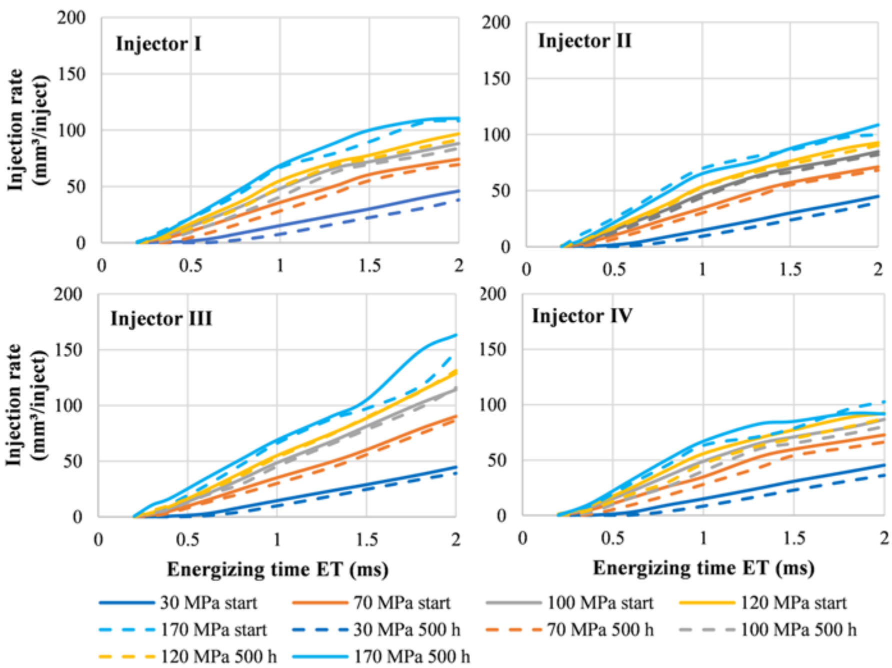


Fig. 3 Injection rate characteristics of the tested electromagnetic injectors

II, III and IV p <0.001. Regarding Injector III, the average value of the coefficient has increased, while in other cases its average value decreased. This has a direct impact on the operation of the injectors, because injection rate control is based on injection timing control with the assumed pressure values and the injection rate obtained thereby. If the injection process itself is changed, it leads to deterioration of the internal combustion engine operation, regardless of whether the obtained flow rate is higher or lower than the reference value. This may cause changes in the achieved engine operating parameters, as well as lead to increased emission of toxic exhaust gas components, predominantly nitrogen oxides and solid particles [19, 21]

In addition, following the performance test, the injectors were tested on the BOSCH EPS 945 test stand. The test has shown that Injector IV neither meets the assumed requirements, nor it is not possible to correct its injection characteristics and thus requires mechanical intervention. As for the other injectors, despite significant changes in their injection rates (Fig. 3), particularly at short injection times, they meet the manufacturer's standards and can therefore be qualified as operational.

An attempt was also made to establish a relationship between the changes in the injector operating characteristics and the surface condition of the interacting injection system components. The literature review shows that the injection rate can be affected not only by the holes but also by wear of the surfaces of the injection system components responsible for the fuel injection process, i.e. the conical surface of the control valve interacting with the valve ball and the surface of the injector needle interacting with the injector tip body [2, 10, 18]. After the performance test, these components were subjected to surface texture examination. Surface examination results obtained for the control valves are shown in Fig. 5 and for the injector needles in Fig. 6.

An analysis of the images in Figures 5 and 6 showing the surface texture of the fuel injector control components demonstrates that no surface degradation took place during the performance test, which means that no component has been damaged or lost its operational ability. Evaluation of this type is usually made by microscopic examination aimed at identifying visible surface defects. The images of the sur-

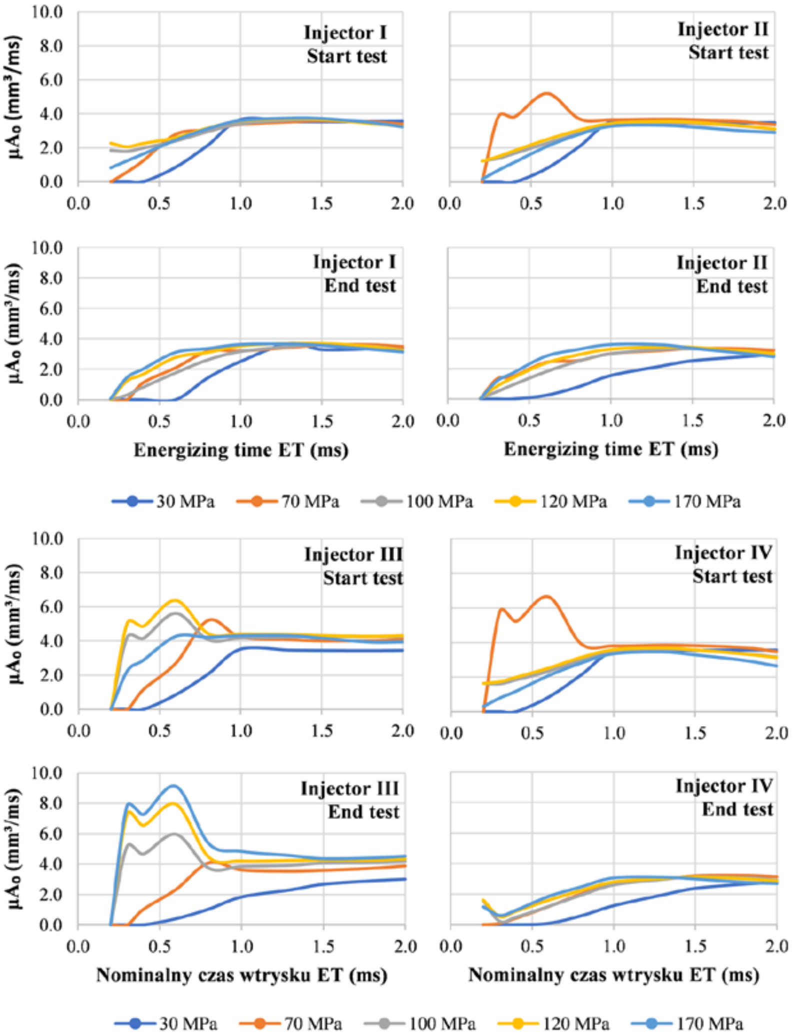


Fig. 4. Characteristics of the μA_0 coefficients μ and A_0 of effective cross-sectional area of the injector

faces of both control valves and injector needles show no visible damage.

An analysis of the surface texture of the control valves was performed in the area of interaction with the valve ball, and obtained results are given in Table 4. An analysis of the Abbott-Firestone curve for the examined surface does not indicate significant degradation of the surface. No increase in the parameters *Vvc* (core void volume) and *Vvv* (valley void volume) is observed with respect to the valve that was not subjected to performance tests. Similarly, the behaviour pattern of the *Sv* parameter (the maximum pit height) does not indicate any damage.

A comparison of the flow characteristics and surface condition of the interacting components reveals that wear processes are not the main cause of changes in the fuel delivery rate. Given the moderate temperature during the tests, the effect of polymeric deposits can be excluded, too. This is evidenced by insignificant changes in the flow coefficient in the range of long injection times. On the other hand, considerable changes in the flow parameters can be observed in the range of short injection times. This means that in the initial stage of operation, the friction of the interacting components changes to a large extent, which affects the dynamics of the pilot valve and injector needle. As a result, small pilot injection quantities can vary greatly during operation and be different for different cylinders.

It should be emphasized that such changes in injection characteristics may be undetected and thus non-compensated for by adaptive algorithms, as the injected fuel quantities are small and do not significantly change the torque. On the other hand, changing the pilot injection significantly affects exhaust emissions. Unfortunately, the observed changes in fuel injection characteristics are quite incidental. The injectors tend to both increase and decrease small fuel quantities. The relationship between fuel pressure and injection rate characteristics is also ambiguous. This proves that the injector operation dynamics depends on many factors. Given the multidirectional nature of the observed changes, it is not possible to develop an injector wear model that could be used to modify the injection strategy.

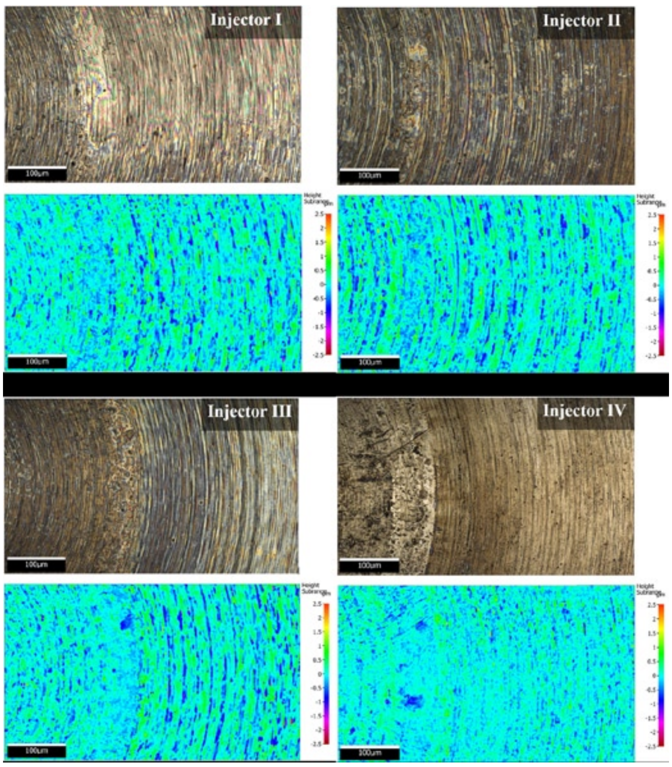


Fig. 5 Images showing the surface texture of control valves

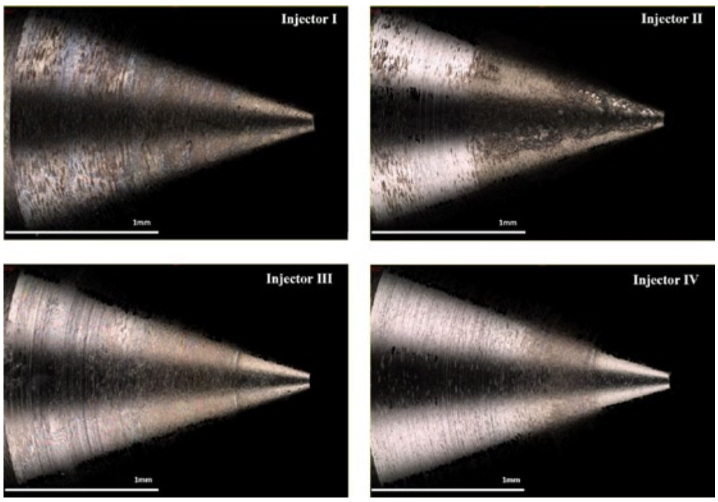


Fig. 6. Images showing the conical surface of nozzle needles

Table 4. Surface texture parameters of the control valves subjected to performance testing

Parameter	New valve	Valve of Injector I	Valve of Injector II	Valve of Injector III	Valve of Injector IV
Vvc (ml/m^2)	0.486	0.354	0.386	0.244	0.291
Vvv (ml/m^2)	0.055	0.040	0.043	0.042	0.0291
Sv (μm)	2.32	1.98	2.04	2.03	2.02

5. Conclusions

In this study, a 500-hour performance test of electromagnetic fuel injectors for a compression-ignition engine was performed. The test was carried out on a test stand by simulating real operating conditions. It was conducted at room temperature to minimize the formation of deposits inside the injectors. Before and after the durability test, injection rate characteristics were determined. Additionally, the changes in the surface texture of the cooperating pairs were evaluated. Based on the results obtained in this study, the following conclusions can be drawn:

- In all tested injectors, statistically significant changes were observed with respect to the effective cross-sectional area of the flow coefficient. However, the direction of these changes differed. For three of the tested injectors, the effective cross-section decreased, but it increased for one of them.
- The examination of the surface of the injector needles and control valves did not reveal any visible damage that could affect the injection process. In addition to that, no significant changes in the texture of their working surfaces were detected.
- Despite the stated lack of wear, the dynamics of injector operation changed to a significant extent. Specifically, there occurred changes in the effective cross-sectional area of flow at short injection times when the valve opening and closing delays play a decisive role, because a stable flow state is not achieved.
- To compensate for the changes in injector operation dynamics, which are important in metering pilot injections, it is necessary to obtain information about every fuel quantity in multi-stage injection. This information can be indirectly obtained by individual measurements of fuel pressure in each of the injectors.

Source of funding

The project/research was financed in the framework of the project Lublin University of Technology-Regional Excellence Initiative, funded by the Polish Ministry of Science and Higher Education (contract no. 030/RID/2018/19).

References

1. Agarwal A K, Singh A P, Maurya R K et al. Combustion characteristics of a common rail direct injection engine using different fuel injection strategies. International Journal of Thermal Sciences 2018; 134: 475-484, https://doi.org/10.1016/j.ijthermalsci.2018.07.001.
2. Asi O. Failure of a diesel engine injector nozzle by cavitation damage. Engineering Failure Analysis 2006; 13(7): 1126-1133, https://doi.org/10.1016/j.engfailanal.2005.07.021.
3. Birgel A, Ladommatos N, Aleiferis P et al. Deposit Formation in the Holes of Diesel Injector Nozzles: A Critical Review. SAE Technical Paper 2008-01-2383 2008, https://doi.org/10.4271/2008-01-2383.
4. Ferrari A, Mittica A, Paolicelli F, Pizzo P. Hydraulic Characterization of Solenoid-actuated Injectors for Diesel Engine Common Rail Systems. Energy Procedia 2016; 101: 878-885, https://doi.org/10.1016/j.egypro.2016.11.111.
5. Gumus M, Sayin C, Canakci M. The impact of fuel injection pressure on the exhaust emissions of a direct injection diesel engine fueled with biodiesel-diesel fuel blends. Fuel 2012; 95: 486-494, https://doi.org/10.1016/j.fuel.2011.11.020.
6. Hofmann O, Han S, Rixen D. Common Rail Diesel Injectors with Nozzle Wear: Modeling and State Estimation 2017, https://doi.org/10.4271/2017-01-0543.
7. Hofmann O, Strauß P, Schuckert S et al. Identification of Aging Effects in Common Rail Diesel Injectors Using Geometric Classifiers and Neural Networks. SAE Technical Paper 2016-01-0813 2016, https://doi.org/10.4271/2016-01-0813.
8. Hunicz J, Beidl C, Knost F et al. Injection Strategy and EGR Optimization on a Viscosity-Improved Vegetable Oil Blend Suitable for Modern Compression Ignition Engines. SAE Int. J. Adv. & Curr. Prac. in Mobility 2020; 3(1): 419-427, https://doi.org/10.4271/2020-01-2141.
9. Hunicz J, Matijošius J, Rimkus A et al. Efficient hydrotreated vegetable oil combustion under partially premixed conditions with heavy exhaust gas recirculation. Fuel 2020; 268: 117350, https://doi.org/10.1016/j.fuel.2020.117350.
10. Ignaciuk P, Gil L. Uszkodzenia kawitacyjne w silnikach spalinowych. Autobusy 2014; 5: 63-65.
11. Jindal S, Nandwana B P, Rathore N S, Vashistha V. Experimental investigation of the effect of compression ratio and injection pressure in a direct injection diesel engine running on Jatropha methyl ester. Applied Thermal Engineering 2010; 30(5): 442-448, https://doi.org/10.1016/j.applthermaleng.2009.10.004.
12. Klonecki W. Elementy statystyki dla inżynierów. Wrocław, O. Wyd. Politechniki Wrocławskiej: 1996.
13. Knefel T. Ocena techniczna wtryskiwaczy Common Rail na podstawie doświadczalnych badań przelewów. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14(1): 42-53.
14. Kurczyński D, Adrian A. Analiza uszkodzeń układu zasilania Common Rail. Logistyka 2014; 6: 6406-6419.
15. Lacey P, Gail S, Kientz J M et al. Fuel Quality and Diesel Injector Deposits. SAE Int. J. Fuels Lubr. 2012; 5(3): 1187-1198, https://doi.org/10.4271/2012-01-1693.
16. Lazarev V, Lomakin G, Lazarev E et al. Optimum Design Of Nozzles' Tribology Systems Of A Diesel Engine Fuel Injector With High Values Of Rail Pressure. WIT Transactions on Ecology and the Environment 2015; 195: 379-389, https://doi.org/10.2495/ESUS150321.
17. Liaquat A M, Masjuki H H, Kalam M A, Rizwanul Fattah I M. Impact of biodiesel blend on injector deposit formation. Energy 2014; 72: 813-823, https://doi.org/10.1016/j.energy.2014.06.006.
18. Ligier K, Oczkowski A. Analiza uszkodzeń wybranych elementów aparatury wtryskowej silników wysokoprężnych. Logistyka 2015; (4): 4473-4479.
19. Mohan B, Yang W, Yu W et al. Numerical investigation on the effects of injection rate shaping on combustion and emission characteristics of biodiesel fueled CI engine. Applied Energy 2015; 160: 737-745, https://doi.org/10.1016/j.apenergy.2015.08.034.
20. Niewczas A, Gil L, Ignaciuk P. Analiza procesów zużycia tribologicznego elementów układu wtryskowego silnika o zapłonie samoczynnym

zasilanego olejami roślinnymi. Zeszyty Naukowe WSEI. Seria Transport i Informatyka 2011; 1(1): 5-21.
21. Nishimura T, Satoh K, Takahashi S, Yokota K. Effects of Fuel Injection Rate on Combustion and Emission in a DI Diesel Engine. SAE Technical Paper 981929 1998, https://doi.org/10.4271/981929.
22. Paykani A, Garcia A, Shahbakhti M et al. Reactivity controlled compression ignition engine: Pathways towards commercial viability. Applied Energy 2021; 282: 116174, https://doi.org/10.1016/j.apenergy.2020.116174.
23. Payri R, Martí-Aldavarí P, Montiel T, Viera A. Influence of aging of a diesel injector on multiple injection strategies. Applied Thermal Engineering 2020; 181: 115891, https://doi.org/10.1016/j.applthermaleng.2020.115891.
24. Payri R, Salvador F J, Gimeno J, Montiel T. Aging of a Multi-Hole Diesel Injector and Its Effect on the Rate of Injection. SAE Int. J. Adv. & Curr. Prac. in Mobility 2020; 2(6): 3347-3355, https://doi.org/10.4271/2020-01-0829.
25. Pehan S, Jerman M S, Kegl M, Kegl B. Biodiesel influence on tribology characteristics of a diesel engine. Fuel 2009; 88(6): 970-979, https://doi.org/10.1016/j.fuel.2008.11.027.
26. Pielecha I, Skowron M, Mazanek A. Evaluation of the injectors operational wear process based on optical fuel spray analysis. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(1): 83-89, https://doi.org/10.17531/ein.2018.1.11.
27. Praca zbiorowa. Zasobnikowe układy wtryskowe Common Rail. Warszawa, WKŁ: 2009.
28. Reitz R D, Ogawa H, Payri R et al. IJER editorial: The future of the internal combustion engine. International Journal of Engine Research 2019; 21(1): 3-10, https://doi.org/10.1177/1468087419877990.
29. Schuckert S, Hofmann O, Wachtmeister G. Experimental investigation into simulated aging effects of common-rail injector nozzles: Influences on injection rate, spray characteristics, and engine performance. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 2019; 234(2-3): 349-362, https://doi.org/10.1177/0954407019855289.
30. Stanik W, Jakóbiec J, Mazanek A. Engine tests for coking and contamination of modern multi-injection injectors of high-pressure fuel supplies compression-ignition engine. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(1): 131-136, https://doi.org/10.17531/ein.2018.1.17.
31. Stępień Z. A study of factors influencing the formation of harmful deposits in the diesel engine injectors. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(3): 331-337, https://doi.org/10.17531/ein.2017.3.3.
32. Stoeck T, Osipowicz T. Analiza uszkodzeń i stopnia zużycia wtryskiwaczy Common Rail Bosch. Autobusy : technika, eksploatacja, systemy transportowe 2013; 14(10): 240-244.
33. Urzędowska W, Stępień Z. Badania wpływu jakości oleju napędowego na uszkodzenia układu wysokociśnieniowego wtrysku paliwa. Nafta-Gaz 2009; R. 65, nr: 789-796.

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Matuszczak M, Żbikowski M, Teodorczyk A. Predictive modelling of turbofan engine components condition using machine and deep learning methods. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 359–370, http://doi.org/10.17531/ein.2021.2.16.

Predictive modelling of turbofan engine components condition using machine and deep learning methods

Indexed by: Web of Science Group

Michał Matuszczak[a,b,*], Mateusz Żbikowski[a], Andrzej Teodorczyk[a]

[a] Institute of Heat Engineering, Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, Nowowiejska 21/25, 00-665 Warsaw, Poland
[b] General Electric Company Polska sp. z. o. o., Al. Krakowska 110/114, 02-256 Warsaw, Poland

Highlights

- 0-10 condition rank of a turbofan life limiting component is predicted.
- Environmental and engine sensors data preceding the condition observation are used.
- Ensemble meta-model of neural networks shown the best performance.
- Support vector machines and gradient boosted models did not match neural nets.
- Linear model demonstrated the worst performance among considered models.

Abstract

The article proposes an approach based on deep and machine learning models to predict a component failure as an enhancement of condition based maintenance scheme of a turbofan engine and reviews currently used prognostics approaches in the aviation industry. Component degradation scale representing its life consumption is proposed and such collected condition data are combined with engines sensors and environmental data. With use of data manipulation techniques, a framework for models training is created and models' hyperparameters obtained through Bayesian optimization. Models predict the continuous variable representing condition based on the input. Best performed model is identified by detemining its score on the holdout set. Deep learning models achieved 0.71 MSE score (ensemble meta-model of neural networks) and outperformed significantly machine learning models with their best score at 1.75. The deep learning models shown their feasibility to predict the component condition within less than 1 unit of the error in the rank scale.

Keywords

reliability, prognostics, deep learning, machine learning, gas turbine, turbofan engine, neural network, condition-based maintenance.

1. Introduction

A modern aircraft's turbofan engine is a complex mechanical system with numerous components that need to be properly maintained to continue its safe and profitable operation. As the components deteriorate they need to be replaced or repaired which drive the engine off wing for often time consuming overhaul [8] and creates a cost burden requiring proper engine fleet management to continue the aircraft operation [18].

Aircraft engine components condition is assessed on recurring inspections and compared to the limits provided by the engine manufacturer which constitute the Instructions for Continued Airworthiness approved and controlled by the regulatory agency in a form of an engine manual [16]. The engine manual limits proposed by the engine manufacturer are based upon understanding of the physics behind the particular wear out scheme and the condition progression until the part cannot be operated any longer and has to be replaced.

With the complexities of loads that parts are exposed to a variety of competing failure modes occuring at different stages of part's age and progresing at different rates comes with significant impact of environmental factors like volcanic activity [12] and air contaminants presence like dust aerosols as seen in a test [6] and in operation [26].

Additionally, an ease of performing a visual on-wing inspection of the hardware depends on its location in the engine and capability of the inspecting crew and its equipement. Thus with all the factors combined the actual confirmation of the part condition is not always feasible.

It is common that engine components wear occurs at different rates and single components compete in being limiting for the engine useful life. Hence a prediction of the current state of the wear of the components becomes a crucial task in the fleet management. With the development of health monitoring systems and on board diagnostics technologies deployment, a significant amount of data has become available for engineers to analyze which enables enhancement of classical condition based maintenance [29].

In the light of the latest research based in the field of predicting components life this paper proposes a data-driven approach for an aviation turbofan engine.

(*) Corresponding author.
E-mail addresses: M. Matuszczak - michal.matuszczak2.dokt@pw.edu.pl, M. Żbikowski - mateusz.zbikowski@pw.edu.pl, A. Teodorczyk - andrzej.teodorczyk@itc.pw.edu.pl

2. Failure prediction methods overview

2.1 Prognostics approaches

There exist numerous examples of attempts to predict the component failure of a part or the entire system depending on the problem at hand, design phase and data available.

In the concept design phase where numerical models are available Ning Baojun et al. proposed a method to incorporate boundary condition uncertainity into the FEA of a turbofan engine combustor to obtain a stochastic life predition [4]. Another approach is presented by Echarda et al. [10] where a SARFAN's aviation engine blade support is analyzed with a variation of geometry, material properties and load variation to computionally capture the life prediction and its probability. These models can be very accurate and deliver useful information about the type design, however a good understanding of the failure mode is necessary.

With available failure data one can apply different predictive methods. In their article Yang et al. explore potential for matching the failure times of an aeronautical equipment components to probability distributions to the outcome of finding that the normal distribution to best reflects the actual life distribution [38]. Whereas some cases show promise of normal distribution use, the others like the subject studied in the other paper by Yang et al. indicate 3 parameter Weibull to best represent failure probability of airborne equipment [37]. These modelling approach enables the engineer to make predictions of the part failure based on the sample of fielded hardware and employing statistical methods in place of finite element computations with a challenge of collecting sufficient amount of well comprehended data.

The other researches focus on the engine health monitoring and fault diagnostics, where engine sensors are used to look for a signal of a deteriorating engine health or a faulty component. Turbofan engine health degradation and prognostics of the remaining useful life (RUL) was deployed by Zaidan et al. [41] with a use of a Bayesian Network Regression. Xiu et al. present an aviation turbofan engine fault diagnosis scheme based on deep belief network (DBN) [36]. The neural network composed of mulitple layers forming restricted Botlzmann machines (RBM) succesfully modeled engine systems and engine sensory data have been fed into the model and corresponding engine fault state have been predicted.

Another deep learning model is researched in a paper by Sina Tayarani-Bathaie et al. and revealed that dynamic neural networks based on multi-layer perceptron (MLP) networks demonstrated promising performance in prediction of a turbofan engine fault [31]. Also, Heimnes in [14] reports a satisfactory results in RUL prediction with a MLP classifier.

In [19] the researchers are introducing useful classifications of the AI-based methodologies used in the aerospace industry for systems health management; (1) knowledge-based, (2) probabilistic and (3) data-driven with authors pointing out towards the growing interest paid by the scientific community to the deep learning methods. Sikorska et al. [30] report successes in the field of prognostics and prediction of RUL by artificial neural networks (ANN) and making them a separate category of RUL prediction models noting their ability to handle noisy data. Pawełczyk et al [25] have recently reported a succesful use of machine learning methods to predict the condition of high pressure compressor in a stationary gas turbine.

A different take on asset failure prediction is presented in the works of Yoon et al. where deep generative models in semi-supervised learning scheme have been implemented to predict estimated time to failure and show that data-driven approaches are alternatives to the physics-driven modelling [40]. In the presented study for the sparse labelled turbofan data the variational autoencoders have delivered great results over the gated recurrent units (GRU) and long short-term memory (LSTM) network architectures.

Among other network architectures deep convolutional neural networks (CNN) have been demonstrated by Babu et al. [3] to be feasible in predicting a capture a non-linear relationship between RUL and sensor data.

Having in mind the mentioned researches in the field of prognostics, the deep learning methods deliver promising results replacing physics based models provided sufficient understanding of the matter is reached as authors demonstrated in number of publications [23].

2.2. Target variable in researches

An important role in prognostics and health management plays a systems health index (HI) as it reflects the system condition and its potential to perform its function throughout the system useful life. The index is widely used concept across researches based in various industries ranging from electronics equipment, through heavy machinery to the aviation industry.

A paper published by Amir et al. has researched a condition-based health index concept where overall health index was calculated based on the individual indicators [1] and used a 10-grade scale differentiating a system condition from good to bad and enabling to categorize the particular system units. In power transformer application a health index ranging from 0 to 1 have been presented by Lata et al. [2] and incorporated a various input relevant to that particular system to establish the resulting index value.

In the case of turbofan engine, a health index based on engine sensor flight by flight data were used to establish and predict a high-pressure compressor deterioration [33,34].

Another interesting way to develop a health index out of turbofan engine sensor readings have been proposed in [5] where a step by step aggregation of the normalized feature values was proposed. In such arrangement a growing health index would cumulate over time of operation and judgements about RUL can be made.

Based on the solid fundaments established by the research community the subject of this paper uses a condition-based health index with 10 grade scale.

3. Problem description

Turbofan engine components are inspected reccurently at least as often as recommended by the engine manufacturer thus providing a valueable condition data. The considered component operates on the condition based maintenance scheme. The participating engines have been monitored for a period from third quarter of 2014 to first quarter of 2020 to obtain one of the hot gas path component data.

The obvious challenge is in the formulation of the life prediction problem. The intent is to determine, based on available information, at what stage of degradation the given component is. A very efficient technique to determine a moment when a given system would fail is RUL estimation. As the authors of the [11] presented, RUL can be determined by use of a degradation characteristic of an aviation engine as input variable to obtain a survival function that later can be used to predict moment of a probable failure. A degradation characteristic is specific to the system and may depend on the physics of a considered wear out mechanism. For a gas turbine it could be an exhaust gas temperature [14] or a compressor recoup pressure [25], both being related to the system wear out and continuous trend of either could be a signature that can be used to judge incoming expiration of useful life.

However, in the researched system, the component wear out, despite progressing with time, is not picked up by engine sensors and thus a trend as such cannot be the degradation characteristic. Also, there exist no spike in any of the sensor readings when the component reaches the condition at which it is desirable to be removed to avoid further costly engine damage and potential impact to the customers operation schedule. Therefore an anomaly detection methods are not available in this case.

Regardless of its lack of visibility in the engine system sensors, the component life is limiting to the entire system. To adress this problem, the authors propose to use component condition data and the engine operation data preceding the inspection at which the condition rank was collected. Then, by the means of data science; conducting data cleaning, feature engineering and feature selection train the models to predict the condition. The expectation behind such an approach is that there might be non-obvious or hard to quantify differences between the engines so that the component in one engine fails at different time that the other. The difference could be operational: frequently fully loaded aircraft, high altitute of an airports used, short climb path, environmental: air aerosols and dusts present, high temperatures at the airport or manufacturing related; tolerances stacking up results in different loads that the component is exposed to. It is expected that, since a turbofan engine is a closed system, these differences can be determined by sensors not directly related to the considered component and those that cannot be otherwise used as a degradation characteristic. Such differences accumulated over the operation time could be resonsible for the condition rank progression at different rate and modern models are anticipated to fit to them.

Due to the data amount, complexity and high non-linearity neural networks are main focus of the research, however machine learning models are used for comparison basis. Once models are developed, it would be possible to use them to monitor the remaining fleet and plan maintenance provided the sensor data would be provided as an input to the models.

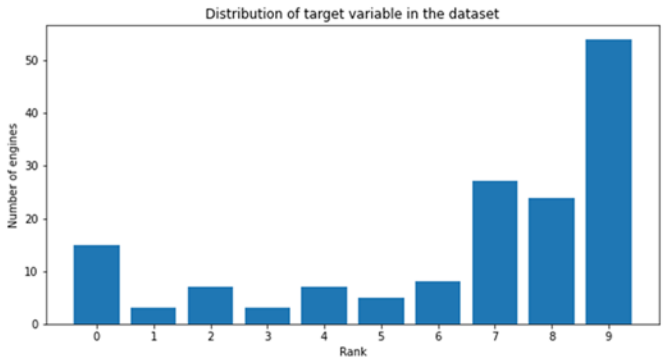


Fig. 1. The number of engines per rank collected during the monitoring program and used as the dataset for this research

Over 150 engines have participated in the monitoring program, running at five different thrust ratings, belonging to 40 different airlines and more importantly operating on different routes across the globe. The engines have been exposed to take-offs and landings in different environmental conditions, altitudes, aircraft loads and runway lenghts, however sharing the same part design. The part condition at the exposure time counted in flight cycles have been recorded. Similarly to authors of [1] a 10-grade scale have been selected to assign meaningful health index, a condtion rank, to the parts based on their actual condition as shown in Table 1. The condition ranks are established based on the inspection limits provided by the engine manufacturer and supported by conclusions from conducting a root cause analysis of this failure mode. In this specific problem, the inspection limits placed in the engine maintenance documentiation have been not sufficient to capture the early progression of the wear and a scale based purely on inspection findings would be highly non-linear. Between the point at which the part exhibits no wear and the point at which first inspection limits for reccuring inspection apply there exist a relatively long period of preceeding damage accumulation that gives away certain symptoms. Upon completed root cause analyses, metallurgical surveys of the components at different damage stages, expert knowledge and numerical simulations the ranks 1-6 have been introduced which improves proportionality of the used scale and makes it more linear. During this procedure limits have been established that enable to assign the rank to inspected hardware. Although, the maintenance documentation enables safe and profitable engine operation, it had to be expanded to be create a proportional scale that can be used in this research to formulate a regression framework. The inspection data have been revisited to assign proper value of the rank per the extended scale as presented in the Table 1. Introduction of new limits that would cause maintenance actions should be carefuly considered as more operation stoppages would be created, driving the aircraft maintenance cost up and are potentially unnecessary. At this stage, authors of this research are trying to study if a model build on such data can deliver results that could be a starting point to reduce the airline maintenance burden by making the findigs at inspection predictable. Nevertheless, as Figure 1 summarizes, the majority of engines labeled are cases requiring replacement and there is a potential class imbalance for a pure classification oriented problem.

As the engine hardware inspection to establish its condition is a recurrent process that needs to be accomodated into the airline maintenance schedule, it puts a time pressure burden with a potential consequence of unplanned delays and it would be beneficial in that regard to obtain a model that could rank the engines prior to obtaining inspection data.

From the perspective of the fleet management such prognostics would enable to plan ahead of time for the replacement hardware delivery and point out to the engines in the fleet needing it first. These are the challenges that authors of this article are trying to adress.

4. Approach

4.1. Dataset creation

Engines are equipped with a number of sensors collecting flight data. Each engine module from front to aft monitors essential operation paramters; pressure, temperature, variable vanes position setting, shafts rotational speeds and fuel flow injected just to name a few. On the top of that, there exist thermodynamics models deployed, validated through testing campaigns, that utilize these readings and

Table 1. 10-grade scale used to assign the health index to the part condition

<table>
<tr><th>Rank</th><th>Condition</th><th>Service limits applicable</th><th>Maintenance action</th></tr>
<tr><td>9</td><td>Not accetable for further operation</td><td>Exceeded</td><td>Engine removal & part replacement</td></tr>
<tr><td>8</td><td>Conditionally acceptable for a short duration</td><td>Allow for operation for short interval</td><td>Increased recurrent inspection frequency on wing</td></tr>
<tr><td>7</td><td>Conditionally acceptable for a long duration</td><td>Allow for operation for long interval</td><td>Recurrent inspection on wing</td></tr>
<tr><td>6</td><td rowspan="5">Wear progression – subsequent expansion of the affected area on the component</td><td rowspan="7">Observed condition is permitted or no specific limits applicable</td><td rowspan="6">Monitoring of the progression on scheduled overhauls when part is exposed</td></tr>
<tr><td>5</td></tr>
<tr><td>4</td></tr>
<tr><td>3</td></tr>
<tr><td>2</td></tr>
<tr><td>1</td><td>Visible wear initiation</td></tr>
<tr><td>0</td><td>No wear confirmed visually</td><td>No action – no wear</td></tr>
</table>

deliver predictions of other useful parameters that are not acquired directly. Additionally, environmental data for arrivals and departure airports are collected with information about ambient temperature, pressure, elevation above sealevel and air aerosols and added to the database. A *Python* programming language with *Keras* [17], *Tensorflow* [34], *Sci-kit learn* [28] and *pandas* [24] libraries are used for data handling and modelling.

Overall the parameters relevant to the engines for which condition-based ranks were established are retrieved from the database and arranged in such a way that every rank at given inspection is preceded by a number of timesteps and the parameters set for each timestep. The strategy to create the dataset is depicted in the Figure 2.

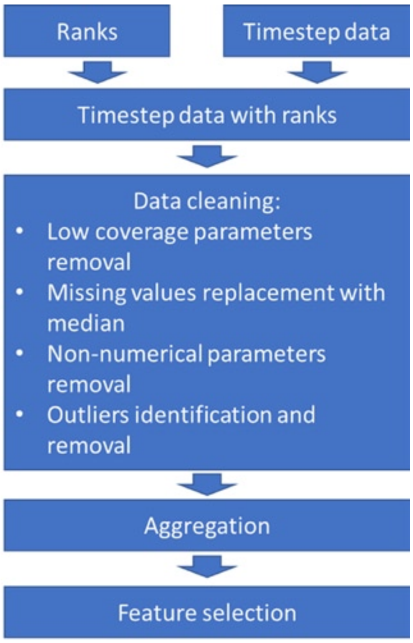


Fig. 2. Dataset creation strategy

In the raw data cleaning process, the parameters having non-numerical values and those not having sufficient coverage over engine operation period are removed. The threshold for lack of coverage is set to be less than 5% of data missing.

Remaining parameters are screened for outlying values, those identified typically come from erroneous sensor readings or faulty data processing and get removed from the set. In an iterative process, all datapoints with standardized score of that parameter, exceeding ± 6σ values are highly suspiscious of being outlying values. Having found such values an investigation has been opened to learn if a sensor malfunctioned, data have been lost or distorted in the migration process or some unexpected event have, in fact, occurred. Upon concluding the investigation, the values were either replaced or removed from the dataset.

The engine's parameters missing values are located and are handled by finding the median value for that particular parameter for the considered engine, then they are filled by that median value. An important consideration is that due to specifics of the aircraft's engine system, each value of parameter should be considered in the missing data management, firstly looking at the data from that engine over time and secondly, if data are too scarce, from the perspective of the sister engine. This minimizes introduction of additional error due to the unknown operational differences.

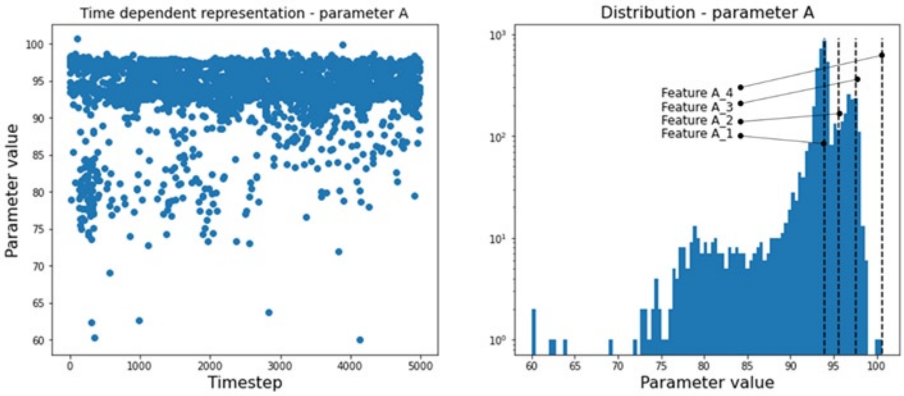


Fig. 3. Feature creation on the example of a single input parameter

4.2. Aggregation and feature selection

To shape the dataset into a problem that can be tackled by machine learning methods the time series data from the sensors are represented by their time independent distributions with the idea depicted in the Figure 3. The values defining the distributions; median, max, 75th percentile value and 95th value are chosen as the new features for the modelling. The selected distribution characteristics come from experimentation with the dataset.

The environmental aerosols data are instead represented by the sum of its departure and arrival values per the flight and accumulated over the total number of flights that engine has completed.

As the engine is a thermodynamic system, a high degree of colinearity is expected between some of its parameters collected during its operation. To adress this issue, a collinerality check is performed within the groups of parameters as shown in the Figure 4. Redundant parameters are identified in this manner that are excluded later from feature creation process.

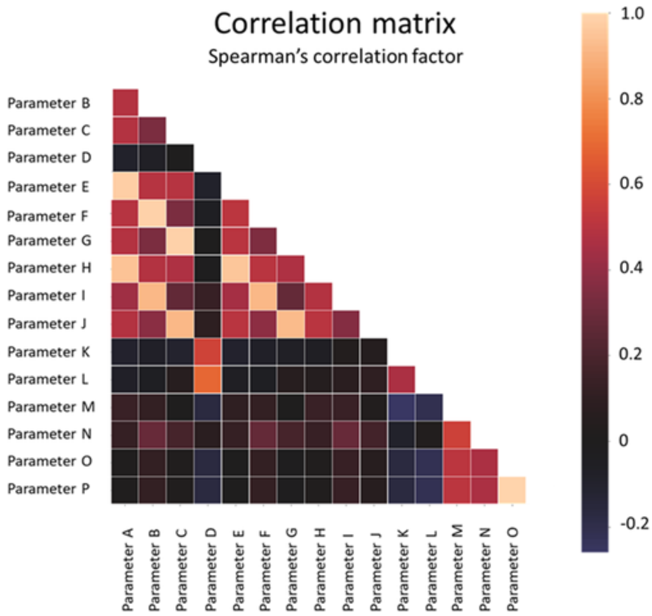


Fig. 4. Correlation matrix

As a final step of feature selection the dataset composed of over 500 features obtained by cleaning and aggregation undergoes a process in which statistically insignificant features are omitted. For that purpose the Boruta algorithm is employed [21]. This procedure limits the number of features to 62 which are later used for developing the best performing model.

4.3. Data transformations

Upon completion of data cleaning and aggregation, the x set is in a form of dataframe of the 62 features by the number of the rows representing the number of the engines and the y are the engine ranks. For the sake of simplicity and having in mind limited number of engines the problem is transformed into a regression problem, where rank is a continuous value from 0 to 9. Additionally, continuous rank is expected to better align with business expectations towards the continuity of the damage progression.

As a next step, the dataset is randomly split into train and validation dataset. The validation dataset is treated as a hold out set and is used eventually to score the models performance against each other. Then, the features are standardized and transformed with *Python scikit-learn* package *StandardScaler* and *PowerTransformer* functions, with the care taken to fitting the functions on the train set, tranforming it and then transforming the validation set, while repeating the procedure feature by feature. The scaling performed by the function follows the equation (1), where x is the value to be scaled:

$$z = (x - \mu) / \sigma \tag{1}$$

μ being a mean value, s is a standard deviation and z is the scaled value.

Additionally, the power transform utilizes the Yeo-Johnson family of equations without the restriction to the values of the variable to be transformed as shown in the equation (2). The input data distribution vary and a transformation to make the distributions more normal is performed. Due to negative values of certain parameters, a simple Box-Cox transformation limited to non-negative values is not feasible. Thus, in the Yeo-Johnson, the λ parameter, representing the transformation parameter, is determined individually for each input feature. In the equation (2), the formulas for λ values at 0 and 2 ensure continuity of the transformation function $\psi(\lambda, y)$ for the entire range y values. The equations for $y \geq 0$ are in fact an equivalent of Box-Cox generalized transformations, whereas the formulas for $y < 0$ enable transformation of negative y values [39]:

$$\psi(\lambda, y) = \begin{cases} \dfrac{\left((y+1)^{\lambda} - 1\right)}{\lambda} & \text{if } \lambda \neq 0,\ y \geq 0 \\ \log(y+1) & \text{if } \lambda = 0,\ y \geq 0 \\ -\dfrac{(-y+1)^{2-\lambda} - 1}{(2-\lambda)} & \text{if } \lambda \neq 2,\ y < 0 \\ -\log(-y+1) & \text{if } \lambda = 2,\ y < 0 \end{cases} \tag{2}$$

4.4. Validation strategy

With the dataset split into train and validation sets, having completed the data cleaning and transformations, a validation strategy for model training, optimization and selection is required.

Hence the train dataset is further used to develop the model, that is to tweak the model and find the best performing hyperparameters on the set. The train set is then often further split into train and test, both complementary subsets of the train set, depending on the need of the specific model. A 7 fold cross-validation (CV) process is used as graphicaly depicted in the Figure 4.1.

As the data become randomly split into k subsets, repeating training over the folds occurs. The model is trained on CV train subset for given set of hyperparameters and scored on CV test subset. In the effect, an average test score from k folds is obtained as shown in the formula (3):

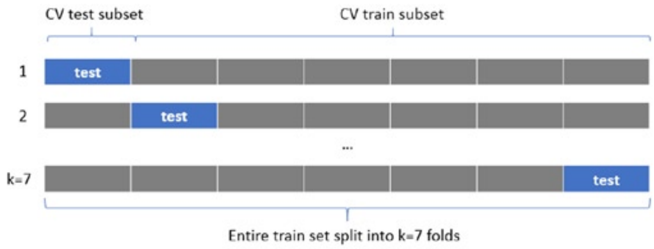


Fig. 4.1. 7 fold cross-validation procedure used in the test

$$CV\ Test\ score = \sum_{i=1}^{k} \frac{test\ score_i}{k} \tag{3}$$

This strategy enables to select the model that performs the best on the train set and has the best average performance while being exposed to the variation present in the train set due to the shuffles made by CV.

The aforementioned validation set is intended to be a hold out set and not used in the model tweaks so that a data leak is avoided and a fair and compenent comparison between the different model possible and to select the one performing best over the specific data. Thus all the comparison scores in this paper are calculated over the validation set via means of multiple further splits into train and test sets with each of the 7 folds of cross-valdation (CV) procedure.

4.4. Hyperparameter optimization strategy

The hyperparameters search is conducted by the means of the Bayesian optimization (BO) [32] where the parameters resulting in the maximum average test score from CV are found. In the Bayesian optimization the objective function $f(x)$ over a dataset is optimized using the benefits of the Bayes' Theorem.

This allows the selection of the most plausible objective function given the prior assumptions regarding the function and hence improve on the performance of the optimization procedure in terms of computational times [7]. In other words, simplifying and applying to the problem at hand, posterior probability of a model M given the evidence (data) E is proportional to the likelihood of E given M multiplied by the prior probability of M (4):

$$P(M \mid E) \propto P(E \mid M) P(M) \tag{4}$$

Instead of *Python scikit-learn* and its *RandomGridSearch* providing the grid search through the hyperparameters,the *bayesopt* package is employed and its implementation of bayesian optimization argument used for every model parameters selection.

4.5. Cost function

As a evaluation score a *mean squared error* (MSE) is calculated as in the equation (5), its used for parameters search in BO and as a mean to compare in between the models. What is more, for the benefit of interpretation ease a R^2 score is calculated however is not used in computations apart from the models comparison:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{5}$$

In the equations (5) and (6) y_i is the ground truth value, also called a target and $\hat{y}_i$ a model prediction:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \mu)^2} \tag{6}$$

5. Models overview

This section describes the models that have been considered for this dataset.

5.1. Linear regression

For the sake of establishing a baseline model for the rank prediction capabilities a linear model is used. The *Ridge* model is used from *Python* package as it incorporates a L2-regularization, called Ridge regression, that helps the model to avoid the overfitting. With the considerate number of features compared to the number of datapoints, the ridge regularization introduces a penalty to the minimization objective by adding the magnitude of sum of square of regression coefficients multiplied by α factor as in the formula (7) where objective is the error to be minimized by the objective function optimization:

$$Loss\ function = \sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \alpha \sum_{j=1}^{p} {\beta_j}^2 \tag{7}$$

5.2. Random Forest and Extremely Randomized Trees

A regressor based on the ensemble of tree predictors is selected for evaluation in the presented problem. The tree predictors are grown over randomly selected inputs and their combinations, offer robustness to outliers and data noise while being fast and additionally due to Law of Large Number they are less prone to overfitting. A random subset of candidate features from the set is used to look for discrimative thresholds via splitting into internal nodes and leafs (external nodes). As the subset is random, the tree shape and the thresholds determining the split cause difference between the estimators which predictions are then averaged out. This becomes a strength of the model as some prediction errors can cancel out. The idea is represented in equation (8):

$$\hat{f}_{rf}(\boldsymbol{x}) = \frac{1}{B}\sum_{b=1}^{B} T_b(\boldsymbol{x}) \tag{8}$$

where B is the number of predictors, T is a tree.

Each of *n_estimators* trees is grown using *max_features* that is used by the algorithm and with tree depth controlled by *max_depth*. Additionally, minimum samples at internal nodes are controlled by *min_sample_split* and at leafs by *min_samples_leaf.*

The *ExtraTreesRegressor* are a variation of the random forest approach that introduces additional randomness as the thresholds at each node are drawn at random and best of them are then used as a splitting rule. Apart from that similar parameters to random forest are defined.

5.3. Support Vector Machines

A non-linear support vector machines regressor with radial basis function kernel is also considered. The support vector machines can be effective in the case where number of features is large compared to the number of samples with the limitation of being memory consuming. From a high-level standpoint and to describe it, a linear example is used. Let the $g(x)$ be a predictor function. If the data is organized in the manner represented in (9) where **x** are input variables vector and the y is the target:

$$(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_l, y_l) \tag{9}$$

the linear predictor function is shown as in (10).

$$g(x) = \boldsymbol{w}, \boldsymbol{x} + b_0 \tag{10}$$

where $\boldsymbol{w}, \boldsymbol{x}$ is dot product of the model weights and the input variables vectors and b_0 an intercept. Transforming this into an optimization problem it takes a form of:

$$minimize \frac{1}{2} \boldsymbol{w}, \boldsymbol{w} \tag{11.1}$$

$$\text{subject to } \begin{cases} y_j - \boldsymbol{w}, \boldsymbol{x_j} - b_0 \le \varepsilon \\ \boldsymbol{w}, \boldsymbol{x_j} + b_0 \le \varepsilon \end{cases} \tag{11.2}$$

In the equations (11) ε represents an error, meaning the weights vector $\boldsymbol{w}$ that results in the solutions lower than error are found. As stated in [35] it is often desirable to have some errors greater than ε and hence the formula is rewritten with introduction of slack variables δ_j and δ_j^* taking the form of these equations (12):

$$minimize \frac{1}{2} \boldsymbol{w}, \boldsymbol{w} + C\sum_{j=1}^{l}\left(\delta_j + \delta_j^*\right) \tag{12.1}$$

$$\text{subject to } \begin{cases} y_j - \boldsymbol{w}, \boldsymbol{x_j} - b_0 \le \varepsilon + \delta_j \\ \boldsymbol{w}, \boldsymbol{x_j} + b_0 - y_j \le \varepsilon + \delta_j^* \\ \delta_j, \delta_j^* \ge 0 \end{cases} \tag{12.2}$$

Upon optimization the first term of the equation is solved just like in (11.1) ensuring weights take low values whereas the second term $C\sum_{j=1}^{l}\left(\delta_j + \delta_j^*\right)$, where l represents the number of observations in the dataset, is known as regularization term and ensures that the optimization problem is feasible. Thus, parameter C offers a trade-off between the model complexity and the error values. Both parameters ε and C are hyperparameters subject to optimization.

5.4. XGBoost

A tree gradient boosting regression model is also researched for feasibility of use on the dataset at hand. This machine learning model has gained popularity due to its performance, speed and scalability. Authors of [9] deliver a very clear description of the algorithm.

The general idea of the model is represented by formula (13), in a dataset $(\boldsymbol{x}_n, y_n)$ composed of n observations and m features in a input vector $\boldsymbol{x}_n$ a tree ensemble model uses K additive functions to predict the target $\hat{y}_i$:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(\boldsymbol{x}_i) \tag{13}$$

Each tree objective function as shown in (14) contains a loss function term which measures the difference between the prediction y_i and the target $\hat{y}_i$ and added regularization term Ω that penalizes the model complexity:

$$Objective = \sum_i loss(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \tag{14}$$

Table 2. Models' hyperparameters overview

	Python package	Model name	Hyperparameters optimized
Linear model	Sklearn.linear_model	*Ridge*	alpha
Random Forest	Sklearn.ensemble	*Random Forrest Regressor*	• max_features • max_depth • min_sample_split • min_samples_leaf • n_estimators
Extremely Randomized Trees	Sklearn.ensemble	*Extra Trees Regressor*	same as in Random Forest
SupportVector Machines	Sklearn.svm	*SVR*	epsilon, C
XGboost	Xgboost	*XGB Regressor*	• max_depth • learning rate • colsample_bylevel • subsample • n_estimators
ANN MLP	Keras/Tensorflow	*Multilayer Perceptron*	• n_layers • n_units per layer • dropout rate • learning rate • test set size • regularization
MLP ensemble	Keras/Tensorflow	*MLP ensemble*	same as in ANN MLP

Loss if a differentiable function that measures the difference between the prediction and the target.

Parameters selected for hyperparameter optimization are as in Table 2.

5.5. Neural networks – multilayer perceptrons (MLP)

Deep neural network is selected as the last type of the model. A multiple hidden layer network, where the input layer takes inputs from the dataset features and then feeds it forwards to a single output neuron predicting the target is built in *Python tensorflow* using *keras* framework.

Let the number of neurons in the layer be m, n the number of samples and k represent the index of the layer. On the very basic level, in the fully connected each neuron in the hidden layer obtains signals vector $\boldsymbol{x}_k$ of m values that represent the input, it gets adjusted by weights assigned to every connection $\boldsymbol{w}_k$ and a bias b_k and then is summed as in equation (15) to create a single output value of the layer v_k. Then an activation function φ is applied on the v_k to obtain the layer output y_k:

$$v_k = \boldsymbol{w}_k \cdot \boldsymbol{x}_k + b_{\boldsymbol{k}} \tag{15}$$

$$\mathrm{v_k} = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} + b_k \tag{16}$$

$$y_k = \varphi(v_k) \tag{17}$$

Then the output becomes input for the next layer neurons and the process repeats until eventually output of the model for a single sample is obtained $\hat{y}_k$. Eventually, the error in the prediction is calculated via loss function by comparison of y_k to the target. Upon the error calculation the back propagation occurs and the error is back propagated via implemented algorithm to adjust all the network weights based on their contribution to the output error.

In the training process the samples are propagated multiple times until the weights are adjusted so that the loss is minimized. The input data is organized in samples and then into smaller batches, which are passed through the model multiple times. In one *epoch* the model has been exposed to all samples in the training set and during one *iteration* the model has adjusted weights to minimize error one batch. In the approach of this research the batch size is set to 1, meaning a model trains on a single randomly selected sample to adjust the weights.

Dropout layers are employed to help prevent the model overfitting, the dropout value is the percentage of neurons in the layer that are randomly excluded from weight adjustment process and do not partake in the output calculation, it is known to contribute to the model robustness. The dropout undergoes hyperparameter optimization. Moreover, a L2 regularization (Ridge) in the first dense layer is turned on, contributing to the objective function with its α value also determined via the optimization process.

As the problem is presented as a regression an activation function is selected to be Parametric Rectified Linear Unit (PReLU). In one of the landmark papers, Kaiming He and others recognized the downfalls of the typically used ReLUok activation function and proposed the alternative which is improvement over Leaky ReLU and demonstrating improvement in image classification error neural network [13]. Thus, the used activation function is as in formula (18):

$$\varphi(v_i) = \begin{cases} v_i, & \mathit{if}\ v_i > 0 \\ \alpha_i v_i, & \mathit{otherwise} \end{cases} \tag{18}$$

It is worth noting that the PReLU behaves like ReLU for positive values of input and the return certain parametric linear output for negative values.

As explained in the chapter 4.3 the train set is used for the model training and optimization leaving the validation set acting as a holdout set. The train set is split in advance into the train and test subsets at random using *StratifiedShuffleSplit* function.

The process repeats k times as the folds of cross-validation enforce model to train and test on a different batch while test set size is maintained. To achieve the perfect balance for this particular dataset, the train to test split ratio is kept as one of the hyperparameters.

Lastly, learning rate is selected as a hyperparameter, meaning the rate at which the weights are adjusted. Importance of this parameter is undoubted as too low values cause inefficient training and too high may cause the model not to converge at all.

Model training, being in the essence finding such model weights, biases and activations, also called parameters that yield the least error, is possible thanks to a gradient descent algorithm [27]. Let $J(\theta)$ be an objective function to be minimized and $\theta \in R$ be the model parameters, by performing the gradient descent, that is updating the parameters in the opposite direction of the gradient of the objective function $\nabla_\theta J(\theta)$ thus following the slope of towards a local minimum. A learning rate η, selected as model hyperparameter in this study, determines the size of the step towards the expected minimum. A popular implementation of this idea, shown in (19), is a stochastic gradient descent (SGD), which enables to calculate the objective function on one sample, instead of all in the batch, that significantly expedites the walk towards the minimum:

$$\theta_t = \theta_{t-1} - \eta \cdot \nabla_\theta J(\theta) \tag{19}$$

Too high value can make the optimization process unstable and prevent the model to converge, too low value can make training process ineffective. There exist numerous optimizers that attempt to improve on it, introducing concepts of momentum to pass over local minima and preventing overshoot due to the overpowering momentum (Nesterov Accelerated Gradient). To better deal with data sparsity an adaptive learning rate algorithm was introduced, Adagrad, that preferentially adjusts learning rates for each parameter and to counteract its downfalls manifesting as monotonically decreasing learning rate Adadelta was proposed. Neural networks trained in the research utilized Adaptive Moment Estimation [20], *ADAM*, that computes adaptive learning rates for each parameter like aforementioned adaptive algorithms but proposing features similar to the concept of momentum. Let the $g = \nabla_\theta J(\theta_t)$ be the gradient and ε be a small term preventing division by zero in the formula (20):

$$\theta_t = \theta_{t-1} - \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon}\hat{m}_t \tag{20}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{21}$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{22}$$

$$m_t = \left(1-\beta_1\right)\cdot g_t + \beta_1 m_{t-1} \tag{23}$$

$$v_t = \left(1-\beta_2\right)\cdot g_t^2 + \beta_2 v_{t-1} \tag{24}$$

where the m_t is a first momentum (23) and the v_t is the second momentum (24) and the β_1, β_2 are decay terms.

5.6. Ensemble

Models collected in an ensemble composed of few best scored neural networks have been explored. In the process of hyperparameter optimization of neural networks, three models with various scores have been obtained. Similar to the concept of the random forest, an ensemble of neural nets can offer an improvement in the overall score as some of the individual model errors can potentially cancel out.

In the study preceding this paper, an ensemble has been created through training meta-model of a similar architecture as single neural network. The meta-model undergoes exactly the same procedure of cross-validated Bayesian hyperparameter optimization with the exception of using the stacked output of the single models as its input and in the prediction is scored with the means of the loss function.

6. Results

Presented results represent the models that have been subjected to hyperparameter optimization described in previous chapters. Both scores R^2 and *MSE* are shown for ease of interpretation, however the *MSE* is selectedfor this regression problem and is used draw conclusions.

The mean squared error score penalizes large errors; as a prediction differs from the true value, the penalty score exhibits quadratic growth. Thus, if used as a loss fucntion in an optimization problem, penalizing large error helps to find model paratmeters that result in minimizing them.

The validation score is calculated over the validation holdout set and the train score represents how model fitted the train set.

As shown in the Figure 5, the best performing model for the specified problem and the data available, has been a neural network meta-model ensemble, achieving *MSE* score of 0.71, that brought 17.4% error decrease from a single best neural network model with a scored at 0.86.

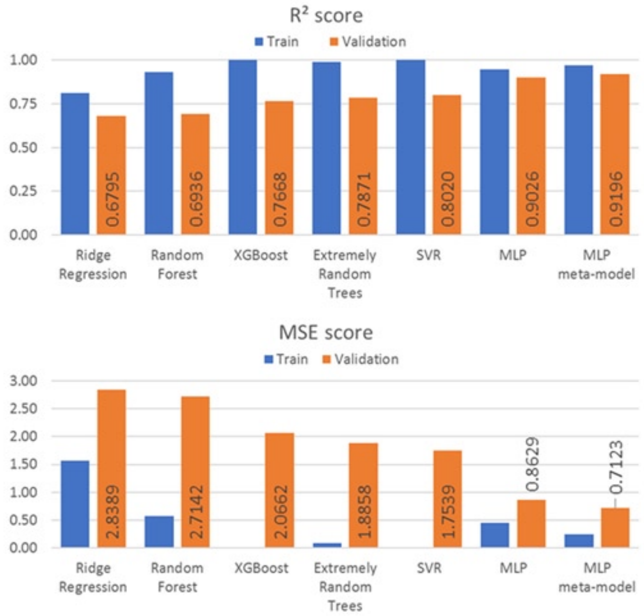


Fig. 5. Results comparison – models' scores. Train and validation series represent model performance on the train and validation sets respectively

The support vector machine regressor model obtained 1.76, that outperformed extremely randomized trees models with a score of 1.88 by a 6.4%. Griadient boosted tree regressor obtained a score of 2.07, random forest model scored 2.71 and ridge regression 2.84.

The difference in error between the score of simple linear model (ridge regression) to the neural net ensemble corresponds to 75% of the linear model score, which justifies the effort invested into deep learning models exploration.

As shown in the Figure 10 even for the best model, there exist outlying residual value in the validation set, which model does not predict well (model underpredicts a 5 distress rank to be little over 3) and increases *MSE* score. Futhermore, a *RMSE* score is also calculated to conclude about the model applicability to the problem at hand.

In addition to the overall models' performance, it has been observed that all researched models have obtained inconsistent score over the ranks as depicted in *RMSE* score plot in Figure 6. Due to scarcity of rank 3 data points, they have not been selected for the validation set via a random selection t*rain_test_split scikit-learn* function. Hence the error values for rank 3 are not available and models ability to predict rank in this range remains not quantified explicitly.

The highest *RMSE* have been produced by SVR (4.01) and linear model (3.66) for rank 1. The lowest *RMSE* values have been achieved by SVR (0.07) and MLP ensemble model (0.10) while predicting rank 0. As demonstrated in the *RMSE* distribution plotted in Figure 6, the most common value is between 1.0 and 1.5.

All studied models have obtained the lowest error while making predictions for rank 7 with similar values scored as quantified by a standard deviation of 0.19 of *RMSE*. Conversely, the greatest inconsistency have been noted for rank 1; MLP-based models scored low error, yet other models have been producing a high error, which contributed to a standard deviation of 1.10 of *RMSE* for this rank.

The ensemble and single neural network models have a better performance for target variable in range from 0-4 (*RMSE* in range from 0.10 to 1.10) and 7-9 (*RMSE* 0.29 to 0.78), than in predicting ranks 5-6 (*RMSE* 1.47 to 2.92). Errors achieved by the MLP based models in this range are the greatest among the considered models followed by XGBoost that have obtained 2.14 *RMSE* over rank 5 and SVR with 1.61 *RMSE* over rank 6.

As a general trend and omitting the exceptionally low errors described earlier, the machine learning models have had higher *RMSE* values for ranks 0-4 (1.01 to 4.01), then error decreases for ranks 5-6 (0.07 to 1.61), becomes the low for all for rank 7 (0.56 to 0.94) and then slightly increases for ranks 8-9 (0.90 - 1.88). This error general trend is different than for earlier discussed MLP-based models.

Some exceptions to this trend have occurred; XGBoost demonstrated greater *RMSE* value for ranks 4-5 (2.14 – 2.42) than for ranks 1-2 (0.94 – 1.62), whereas other machine learning models *RMSE* were in a range of 0.07 – 1.61.

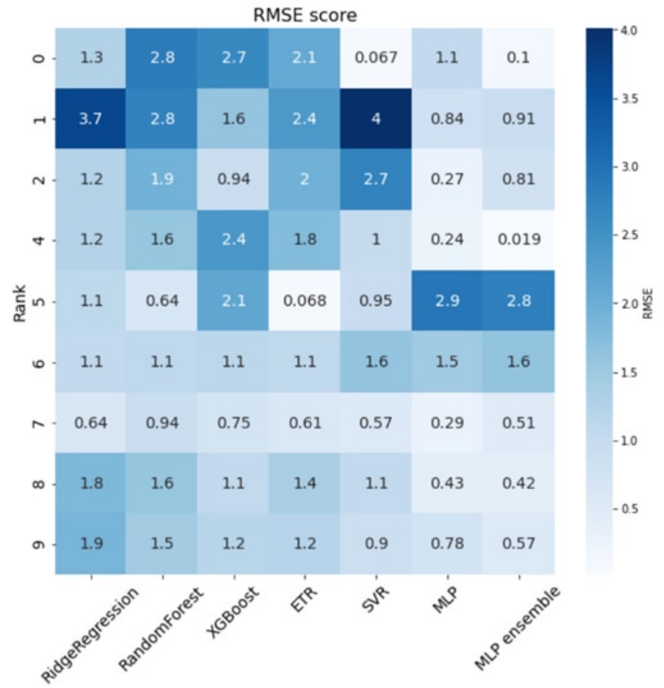


Fig. 6. RMSE score per rank (lower value = less error)

In the tree ensemble based models group; random forest and extremely randomized trees, the latter have, in general, predicted with lower *RMSE* values and offered an improvement in minimum and maximum values. The minimum and maximum values have improved from 0.64 and 2.85 to 0.07 and 2.38, respectively.

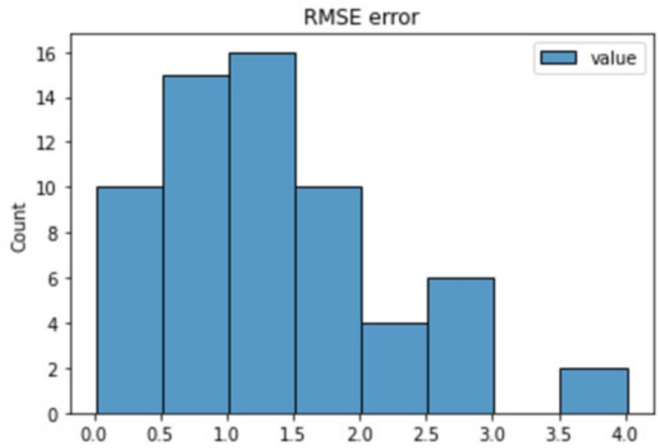


Fig. 7. Distribution of RMSE errors calculated per rank for every model using validation set

Gradient boosted trees model, XGBoost, has surpassed the ETR and random forest models by achieving lower *RMSE* value for ranks 1-2 (XGBoost: 0.94 – 1.62, tree ensemble models: 1.94 – 2.75), however predicted with greater error for ranks 4-5 (XGBoost: 2.14 – 2.42, tree ensemble models: 0.07 – 1.78) and offered some improvement for ranks 8-9 (XGBoost: 1.05 – 1.20, tree ensemble models: 1.21 – 1.59).

SVR *RMSE* values have been low for rank 0 (0.07) and rank 9 (0.9) and comparable to those of MLP ensemble model errors (rank 0: 0.1, rank 9: 0.57). Unfortunately, its prediction error inconsistency through other ranks have been relatively high (*RMSE* 0.6 – 4.01).

Based on the plot in Figure 6 the MLP-based models can make a prediction of low and high ranks with the least error.

The described trends do not correlate with the distribution of the ranks in the training set, training set distribution is similar to that of the entire dataset shown in Figure 7. The data points with rank 9 are most frequent, ranks 8 and 7 occur more rarely and the other ranks data is rather limited. Either of the earlier described trends can be explicitly explained by the distribution of the target variable in the training set.

As the MLP ensemble model predicts with the least error, it is selected as a reference point and the differences in *RMSE* of the others models to the ensemble are calculated and summarized in the plot in Figure 8. The negative difference values, coloured by the shades of red are cases where models have performance debit to the MLP ensemble and conversely, positive values and shades of green show where other models predicted with lower error.

The single MLP model have had a *RMSE* greatest differences for rank 0 (-0.98) and rank 4 (-0.22). MLP ensemble greatly improved error in predicting rank 0. Otherwise, the differences in majority of ranks are between -0.22 and 0.22 values and can be considered similar. An exception to this observation is a rank 2 where single MLP predicted with lower error and the differences was 0.45. Although, there have been ranks where single MLP outperformed the meta-model, the opposite situation has been as frequent and due to the lower overall prediction error, the ensemble model has shown a better performance.

The ensemble meta-model has brought improvement in prediction error it is lower and upper ranges of the target variable. The other models have had, in general, up to -3.10 difference for ranks 0-4 and up to -1.40 difference for ranks 8-9. The models have been within -0.42 to 0.22 in difference to the ensemble for rank 7, with SVR having the least difference (-0.05) and random forest having the greatest

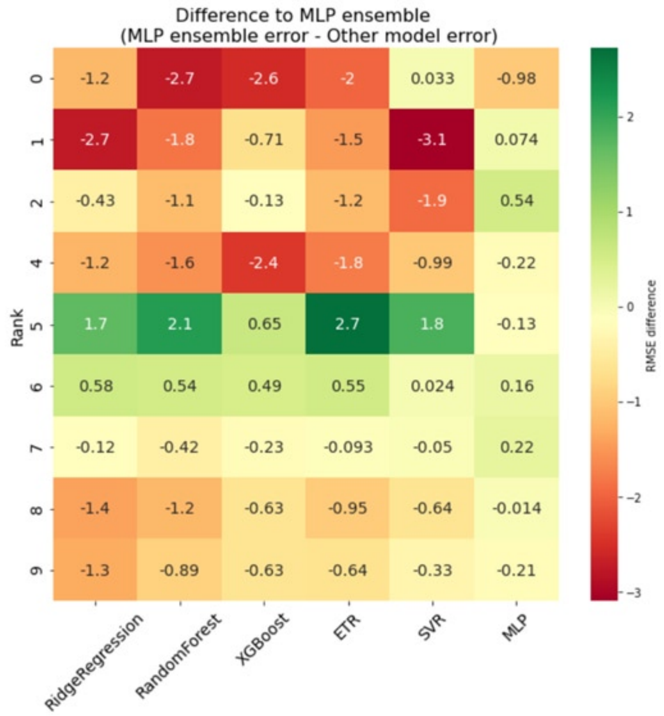


Fig. 8. Difference in error with respect to MLP ensemble model

difference (-0.42). As can be observed, these models outperformed the ensemble in predicting ranks 5-6 with difference up to 2.72 (ETR).

Residual values calculated as a difference between the true and predicted values have been calculated for each model over the train and validation sets and demonstrated for selected models in Figure 9. Non-linear models representing different algorithms families have been chosen: ETR, SVR, XGBoost and MLP.

SVR and XGBoost models have overfitted to the train set, as all prediction values line up closely with their corresponding true values with little residual error, while the validation set residuals are significantly greater. In this particular application, MLP and ETR seem to be less prone to this behaviour and greater train set residuals are visible.

Studied models have also been predicting different outlying values, however due to the noise in the residual values have been hard to interpret. The following observations regarding outlying values have been noted:

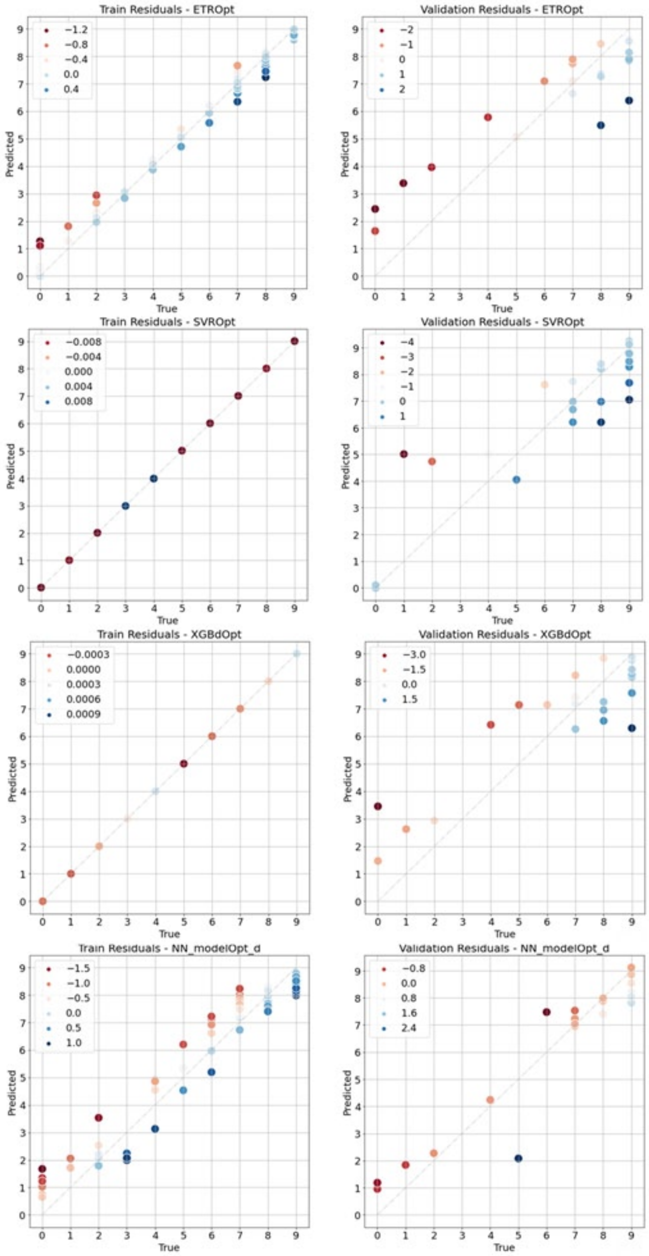


Fig. 9. Residuals plots for selected models

- ETR predictions have the most consistent absolute residual values in the group considered and there are no clear outlying values in the prediction.
- SVR model predicted two outlying values (overpredicted rank 1 and 2).
- XGBoost model residuals are noisy with perhaps one outlying value (overpredicted rank 0).
- MLP predicted one outlying value (underpredicted rank 5).
- SVR, XGBoost and MLP do not predict the same outlying values.

Furthermore, a tendency in over and underprediction have been analysed; XGBoost tends to overpredict the lower ranks and underpredict higher ranks. Similar, however less pronounced, trend is exhibited by ETR. The bull's eye prediction of SVR for rank 0 seems to be an exception and if treated as an outlier, its prediction residual error trend would become similar.

The MLP model is the least noisy in the considered group and does not show a residual error trend exhibited by the other models. What is more, the meta-model ensemble residuals depicted in Figure 10 are similar to the single MLP in lack of the residuals trend and also predict the same outlying value. This explains why ensemble model shares similar performance for rank 5 and demonstrates the ensemble model have not improved the capability to predict this value.

7. Conclusions

Based on results one can observe that certain models have performed better than the others over the given dataset. The promising results presented in the paper align with the recent conclusions of the research community regarding deep learning models applications.

The specifics of the problem have shown that a simple linear model, although useful to certain degree, can be surpassed in performance by more complex architectures. What is more, the superiority of the ensemble model over single neural net model is further confirmed and found in the referenced literatures researchers insights. Additionaly, the neural nets outperformed tree based models and support vector machines. As illustrated in the results, all models have a tendency to overfit to the train set, despite the counter measures taken, however boosted trees, extremely random trees and support vector machines have gravitated towards overfitting more than the others. It might be noted, that the models that have had the lowest difference between train score and validation score are deep learning models. In the effect, their highest validation scores on this dataset could be attributed to their ability to generlize the best and learn without overfitting to the training set.

The best model residuals demonstrate fairly consistent error in continously predicting conditions ranks across the scale and hence it is concluded that it could be satisfactory used for the problem at hand. Translating the *MSE* 0.71 to *RMSE* returns value of 0.84, which, from the forecast perspective, enables to predict ranks with error lower than one condition rank in the scale. Such perspective places the deep learning models considered in this paper as an adequate candidates for the business use, however leaves a room for improvement for future studies for the research community.

The obtained results demonstrate that a neural network model build on the gathered data can predict the rank with average error less than one unit of the rank scale. Although certain models error has not been consistent over the enitre rank scale, a potential business application could benefit by a prediction by few models, keeping in mind their different performance in different rank scale ranges. As a conclusion it may be underlined, that proper data collection and ranking the collected inspection data is a relatively long processes, that is greatly expedited by using established inspection procedures and their findings.

An important challenge has become a selection of a proper rank scale, which should ensure proportionality to formulate a valid regression framework. In the specific example, the existing data based on the engine service limits had to be expanded by introduciton of ranks that represented early wear stages and would normally be omitted per the existing inspection requirements as being acceptable to operate with. Additional ranks required revisiting the collected inspection data and proper re-assignment based on the established scale. The development of the scale required a study of the failure mode, conducting destructive tests, application of material knowledge and involvment of industry experts and wihtout this preceding step further research would not be possible.

In the data collection process, a strong bias towards having the majority of data points composed of worn out parts or parts near the end of its useful life have been observed. This is due to the fact, that in the aviation industry, the airlines tend to maximize the time that aircraft is in operation and stopagges due to the inspections and repairs are additional financial burden. Therefore components near its service limits or requiring recurrent inspections of increased frequency are removed earlier. This data is most widely accesible and shared with the engine manufacturer, which explains the bias in the dataset. On the other hand, due to some unexpected events, i.e. foreign object damage to the engine, the component becomes exposed before the wear process is initiated and the dataset has more data points of this stage than few of the subsequent ranks. The least available data are from the early progression stage of the wear from initiation point to the moment of first service limits apply. This is explained by the fact, that such data is considered acceptable per the inspectors and typically not captured in the inspection process as it presents hardware condition that will continue to operate for a significant time between the wear out. This mindset is a challenge for implementation of a data collection process that enables building a high fidelity prediction model, where a model should be trained with a balanced dataset to predict over the entire range of the target variable with an acceptably low error. With such limitation, ranking scale selection process may become a trade off between having sufficiently many grades to capture the physics and number of data points per each rank for the model to be able to fit to it. As a conclusion from this research, implementation of a data collection scheme expanding the scope of the current inspection data would enable further development of such models. However, it should be noted, that a potential data collection processes to keep the models up to date can be done without the modification of the inspection limits and done post inspection by the engine manufacturer. This approach would help to reduce the maintenance cost by providing a way to monitor fleet's health and manage the maintenance without creating additional opeartion stoppages.

Using the model, a prediction for every turbofan engine condition in the fleet can be obtained easily and updating the prediction regularly with the new input data can provide useful information about the progression of the wear and change in the fleet's health. Information about the rank could enable to schedule maintenance and set expectations regarding the condition once engine is visually inspected on-wing. The information available ahead of time can enable a prioritization of engine repairs and ordering replacement hardware. Presented study demonstrates that use of such data can deliver a valuable solution to the industry with relatively low investment of time and resources using the latest developments in deep and machine learning. In the nearest perspective, models might not be feasible to replace the on-wings inspection, but can reduce an inspection burden by making its outcomes more manageable and predictable. Safety has always been a number one factor in the aviation industry and the most likely application of such models is expected in the fleet health monitoring and maintenance management rather than direct replacement of well established inspection processes.

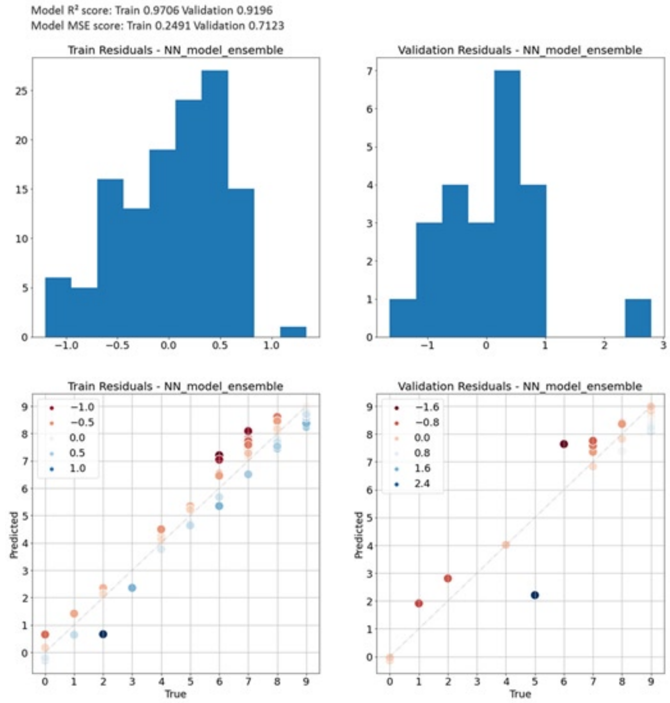


Fig. 10. Ensemble meta-model residuals

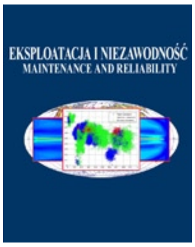
8. Next steps

Authors of the article recognize the promising results obtained by the scientific community using recurring neural networks architectures in similarly stated problems, the demonstrated performance of deep Bayesian networks and the advantages of combining the efficiency of semi-supervised learning variational autoencoders with deep Bayesian network models on sparsely labelled data typically encountered in the aviation industry, thus wish to try these methods to further research this particular problem.

References

1. Amir M D M, Muttalib E S A., Health index assessment of aged oil-filled ring main units. IEEE 8th International Power Engineering and Optimization Conference, 24-25 March 2014, https://doi.org/10.1109/PEOCO.2014.6814452.
2. Azipurua J I, Stewart B G, McArthur S D J, Lambert B, Cross J G, Catterson V M., Improved power transformer condition monitoring under uncertainty through soft computing and probabilistic health index., Applied Soft Computing 2019; 85, https://doi.org/10.1016/j.asoc.2019.105530.
3. Babu G S, Zhao P, Li X L., Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life., Database Systems for Advanced Applications. DASFAA 2016. Lecture Notes in Computer Science 9642, https://doi.org/10.1007/978-3-319-32025-0_14.
4. Baojun N, Mei Y, Xingjian S, Peng W. Random FEA and reliability analysis for combustor case., 2017 Prognostics and System Health Management Conference (PHM-Harbin), Harbin 2017: 1-5, https://doi.org/10.1109/PHM.2017.8079268.
5. Bektas O, Jones J A, Sankararaman S, Roychoudhury I, Goebel K., A neural network filtering approach for similarity-based remaining useful life estimation., The International Journal of Advanced Manufacturing Technology 2019; 101: 87-103, https://doi.org/10.1007/s00170-018-2874-0.
6. Bojdo N, Ellis M, Filippone A, Jones M, Pawley A. Particle-Vane Interaction Probability in Gas Turbine Engines., Journal of Turbomachinery 2019, 141(9), https://doi.org/10.1115/1.4043953.
7. Brochu E, Cora V M, de Freitas N. A Tutorial on Bayesian Optimization of Expensive Cost Functions with Application to Active User Modeling and Hierarchical Reinforcement Learning 2010, https://arxiv.org/abs/1012.2599.

8. Cerdeiera J O, Lopes I C, Silva E C., Scheduling the Repairment of Aircrafts' Engines., 2017 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO), Prague, 2017: 259-267, https://doi.org/10.1109/ICCAIRO.2017.56.
9. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 2016: 785-794, https://doi.org/10.1145/2939672.2939785.
10. Echarda B, Gaytona N, Bignonnet A. A reliability analysis method for fatigue design. International Journal of Fatigue 2014: 59: 292-300, https://doi.org/10.1016/j.ijfatigue.2013.08.004.
11. Gao Z, Li Jiwu, Wang R. Prognostics uncertainty reduction by right-time prediction of remaining useful life based on hidden Markov model and proportional hazard model. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(1): 154-164, https://doi.org/10.17531/ein.2021.1.16.
12. Guffanti M, Tupper A. Chapter 4 - Volcanic Ash Hazards and Aviation Risk. Volcanic Hazards Risks and Disasters 2015: 87-108, https://doi.org/10.1016/B978-0-12-396453-3.00004-6.
13. He K, Zhang X, Ren S, Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago 2015: 1026-1034, https://doi.org/10.1109/ICCV.2015.123.
14. Heimes F. Recurrent neural networks for remaining useful life estimation. 2008 International Conference on Prognostics and Health Management 2008: 1-6, https://doi.org/10.1109/PHM.2008.4711422.
15. Illustration, https://res.mdpi.com/information/information-10-00113/article_deploy/html/images/information-10-00113-g003.png (accessed 10 October 2020).
16. Instructions for Continued Airworthiness, Code of Federal Regulations 14 CFR § 33.4 (1980).
17. Keras documentation, https://keras.io/api/ (accessed 10 April 2020).
18. Khan N, Manarvi I A. Identification of delay factors in C-130 aircraft overhaul and finding solutions through data analysis. 2011 Aerospace Conference, Big Sky, MT, 2011: 1-8, https://doi.org/10.1109/AERO.2011.5747591.
19. Khan S, Yairi T. A review on the application of deep learning in system health management. Mechanical Systems and Signal Processing 2018; 107: 241-265, https://doi.org/10.1016/j.ymssp.2017.11.024.
20. Kingma D P, Ba J. ADAM: A Method for Stochastic Optimization. 3rd International Conference for Learning Representations 2015: 1-13, https://arxiv.org/abs/1412.6980.
21. Kursa M B, Rudnicki W R. Feature Selection with the Boruta Package. Journal of Statistical Software 2010; 36(11), http://dx.doi.org/10.18637/jss.v036.i11.
22. Li Y X, Shi J, Gong W, Zhang M. An ensemble model for engineered systems prognostics combining health index synthesis approach and particle filtering. Quality and Reliability Engineering International 2017; 33(8): 2711-2725, https://doi.org/10.1002/qre.2229.
23. Malik K, Zbikowski M, Teodorczyk A. Detonation cell size model based on deep neural network for hydrogen, methane and propane mixtures with air and oxygen. Nuclear Engineering and Technology 2019; 51(2): 424-431, https://doi.org/10.1016/j.net.2018.11.004.
24. Pandas documenation, https://pandas.pydata.org/pandas-docs/stable/ (accessed 20 September 2020).
25. Pawełczyk M, Fulara S, Sepe M, De Luca A, Badora M. Industrial gas turbine operating parameters monitoring and data-driven prediction. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(3): 391-399, https://doi.org/10.17531/ein.2020.3.2.
26. Przysowa R, Gawron B, Kulaszka A, Placha-Hetman K. Polish experience from the operation of helicopters under harsh conditions. Journal of Konbin 2018; 48(1): 263-299, https://doi.org/10.2478/jok-2018-0056.
27. Ruder S. An overview of gradient descent optimization algorithms, 2017, https://arxiv.org/abs/1609.04747.
28. Scikit-learn documentation, https://scikit-learn.org/stable/index.html (accessed 20 September 2020).
29. Shin J H, Jun H B. On condition based maintenance policy. Journal of Computational Design and Engineering 2015; 2(2): 119-127, https://doi.org/10.1016/j.jcde.2014.12.006.
30. Sikorska J Z, Hodkiewicz M, Ma L. Prognostic modelling options for remaining useful life estimation by industry. Mechanical Systems and Signal Processing 2011; 25(5): 1803-1836, https://doi.org/10.1016/j.ymssp.2010.11.018.
31. Sina Tayarani-Bathaie S, Sadough Vanini Z N, Khorasani K. Dynamic neural network-based fault diagnosis of gas turbine engines. Neurocomputing 2014; 125: 153-165, https://doi.org/10.1016/j.neucom.2012.06.050.
32. Snoek J, Larochelle H, Adams R P. Practical Bayesian Optimization of Machine Learning Algorithms. Advances in Neural Information Processing Systems 2012; 25, https://arxiv.org/abs/1206.2944.
33. Sun J, Zuo H, Wang W, Pecht M G. Application of a state space modeling technique to system prognostics based on a health index for condition based maintenance. Mechanical Systems and Signal Processing 2012; 28: 585-596, https://doi.org/10.1016/j.ymssp.2011.09.029.
34. Tensorflow documentation, https://www.tensorflow.org/ (accessed 9 May 2020).
35. Wu S, Akbarov A. Support Vector Regression for Warranty Claim Forecasting. European Journal of Operational Research 2011; 213(1): 196-204, https://doi.org/10.1016/j.ejor.2011.03.009.
36. Xu J, Liu X, Wang B, Lin J. Deep Belief Network-Based Gas Path Fault Diagnosis for Turbofan Engines. IEEE Access 2017; 7: 170333-170342, https://doi.org/10.1109/ACCESS.2019.2953048.
37. Yang Y, Ding Y,and Zhao Z. Fault distribution analysis of airborne equipment based on probability plot. 3rd IEEE International Conference on Control Science and Systems Engineering 2017: 239-242, https://doi.org/10.1109/CCSSE.2017.8087931.
38. Yang Y, Guo F. Reliability Analysis of Aero-Equipment Components Life Based on Normal Distribution Model. IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC) 2018, Chongqing, China, 2018: 1070-1074, https://doi.org/10.1109/ITOEC.2018.8740448.
39. Yeo I K, Johnson R A. A new family of power transformations to improve normality or symmetry. Biometrika 2000; 87(4): 954-959, https://doi.org/10.1093/biomet/87.4.954.
40. Yoon A S, Lee T, Lim Y, Jung D, Kang P, Kim D, Park K, Choi Y. Semi-supervised Learning with Deep Generative Models for Asset Failure Prediction. KDD17 Workshop on Machine Learning for Prognostics and Health Management 2017, Canada, https://arxiv.org/abs/1709.00845.
41. Zaidan M A, Harrison R F, Mills A R, Fleming P J. Bayesian hierarchical models for aerospace gas turbine engine prognostics. Expert Systems with Applications 2015; 42(1): 539-553, https://doi.org/10.1016/j.eswa.2014.08.007.



Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Zhu Q, Sun W, Zhoua Y, Gao C. A tool wear condition monitoring approach for end milling based on numerical simulation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 371–380, http://doi.org/10.17531/ein.2021.2.17.

A tool wear condition monitoring approach for end milling based on numerical simulation

Indexed by: Web of Science Group

Qinsong Zhu [a], Weifang Sun [a], Yuqing Zhou [a,*], Chen Gao [b,*]

[a] College of Mechanical and Electrical Engineering, Wenzhou University, Wenzhou, China
[b] School of Mechatronics and Transportation, Jiaxing Nanyang Polytechnic Institute, Jiaxing, China

Highlights

- A numerical simulation model is proposed to overcome sample missing and insufficiency.
- Model parameters are optimized by orthogonal experiment and KL divergence.
- The optimized model provide effectively missing samples and expand sample size.
- Experimental results show the proposed method improves notably the performance of TCM.

Abstract

As an important research area of modern manufacturing, tool condition monitoring (TCM) has attracted much attention, especially artificial intelligence (AI)- based TCM method. However, the training samples obtained in practical experiments have the problem of sample missing and sample insufficiency. A numerical simulation- based TCM method is proposed to solve the above problem. First, a numerical model based on Johnson-Cook model is established, and the model parameters are optimized through orthogonal experiment technology, in which the KL divergence and cosine similarity are used as the evaluation indexes. Second, samples under various tool wear categories are obtained by the optimized numerical model above to provide missing samples not present in the practical experiments and expand sample size. The effectiveness of the proposed method is verified by its application in end milling TCM experiments. The results indicate the classification accuracies of four classifiers (SVM, RF, DT, and GRNN) can be improved significantly by the proposed TCM method.

Keywords

tool wear, sample missing, sample insufficiency, numerical simulation, cutting force.

1. Introduction

Computer numerical control (CNC) milling machines, which stable and efficient operation can produce huge economic value, are the most widely used automatic production equipment in modern manufacturing industry. The milling tool is the most critical and vulnerable part in the milling process, its wear state affect directly the surface quality of the machined parts and the normal operation of the machine tool [20, 21, 37]. Therefore, it is particularly important to develop an accurate tool condition monitoring (TCM) method.

The above-discussed issue has been addressed in the past few years by developing two general types of TCM methods, direct TCM method and indirect TCM method. The direct TCM method is often seldom adopted because it is greatly affected by the machining environment, such as light, cutting chips, and cutting fluid [49]. In contrast, the indirect TCM method employs certain artificial intelligence (AI) classifier to predict the wear state through collecting sensor signals associated with the tool wear state [33], such as cutting force [15, 50], vibration [6], acoustic emission (AE) [44], and motor current [47], sound [19, 46] signals. Recently, with the development of artificial intelligence (AI) algorithms, more and more scholars have applied AI algorithms in TCM, including support vector machine (SVM) [6,18], random forest (RF) [24, 32, 41], decision tree (DT) [3, 26], artificial neural network (ANN) [1, 9, 12, 22, 23, 28, 34]. However, while these AI methods have yielded encouraging achievements in TCM applications, achieving good wear state prediction performance using these methods relies heavily on large datasets of monitoring signals that are associated with all possible tool wear conditions for model training [14, 45], which is costly and time-consuming for machining processes under different cutting conditions. Although SVMs are suitable for model training with small datasets, they are invalid for sample missing as samples associated with some tool wear conditions are often missing due to the complex conditions encountered in the machining process.

Therefore, a low-cost and easy-to-implement method is needed to solve the problem of sample missing and sample insufficiency. In recent years, the numerical simulation technology was promoted by the improvement of computer technology, more and more researchers have begun to pay attention to this technology [16, 27, 40, 43]. For example, Xiang et al. [42] proposed a personalized diagnosis method of shaft based on numerical simulation, combined with wavelet packet transform (WPT) and SVM model to realize the diagnosis of different shaft faults. Gao et al. [7] solved the problem of missing and insufficient samples of bearing faults by combining finite element

(*) Corresponding author.
E-mail addresses: Q. Zhu - 1053876272@qq.com, W. Sun - vincent_suen@126.com, Y. Zhoua - zhouyq@wzu.edu.cn, C. Gao - gaochen_1993@163.com

simulation (FEM) and Generative adversarial networks (GANs), and provided complete training samples for AI models.

The metal cutting process can be understood as a process in which the tool and the workpiece move and collide with each other. The actual machining process can be simulated by establishing models and mathematical expressions. At present, there are a lot of commercial software (such as Deform, AdvantEdge, Abaqus.) in the market that encapsulate the above process in the software to bring convenience to users. The rich functions of these softwares provide the potentials to simulate physical signal corresponding to tool wear state, which can overcome the problem of sample missing and insufficient. Therefore, a novel tool wear condition monitoring method based on numerical simulation is proposed in this paper, and the remainder of this paper is organized as follows. Section 2 introduces the basic working principles of the proposed method, including numerical simulation based on Johnson-Cook (J-C) constitutive model, parameter optimization of the J-C model, and the framework of the proposed method. Experimental investigations with end milling TCM are given in Section 3. Section 4 analyzes the performance of the proposed method. Finally, conclusions are given in Section 5.

2. Proposed method

2.1. Numerical simulation based on J-C model

The essence of the cutting process is that the workpiece material from elastic deformation to the material yield point under the action of external forces, which causes the plastic deformation of workpiece and finally to the process of fracture. In this process, the tool contact and rub against with the workpiece surface and chips to produce wear, cutting force and heat will also be generated between the tool and workpiece. In cutting simulation, material constitutive models are employed to describe this complex process, and the J-C model is often used because it can describe the behavior of high temperature, high strain, and high strain rate in the cutting process. The formula of the J-C model is as follows [36]:

$$\sigma = \left(A + B\varepsilon^{\mathrm{n}}\right)\left(1 + C\ln\left(\frac{\varepsilon'}{\varepsilon_0'}\right)\right)\left(1 - \left(\frac{T - T_0}{T_{melt} - T_0}\right)^{\mathrm{m}}\right) \tag{1}$$

where A (MPa) is the initial yield stress, B (MPa) is the strain hardening coefficient, ε is the the plastic strain, n is the stain hardening exponent, C is the strain rate sensitivity coefficient, ε' and ε_0' are the plastic strain rate (s^{-1}) and reference plastic strain rate (s^{-1}), respectively, T is the deformation temperature of the workpiece (°C), T_0 is room temperature (20°C), T_{melt} is the melting temperature of the material (°C), and m is the temperature softening exponent [13]. the three individual terms separately enclosed within parentheses on the right side of formula (1), respectively, represent the strain strengthening effect of the material, the relationship between σ and the natural logarithm of the relative strain rate, and the exponential relationship between σ and temperature.

Because of the rich material library and the specialized cutting module in software DEFORM, it has attracted the attention of many researchers to the software. Shao et al. [35] adopted DEFORM to study the thermodynamic constitutive equation of Ti-6Al-4V and predicted the tool wear depth. Klocke et al. [17] utilized DEFORM to inverse the constitutive equations and damage criteria of AISI 1045 and Inconel 718, and verified the effectiveness of the method by comparing simulation results with experimental results. Thus, the software DEFORM is used in this paper to simulate the end milling process and obtain the missing wear samples.

2.2. Parameter optimization

The benchmark value of five parameters, A, B, n, C, m, in the J-C model with certain workpiece material can be obtained from split Hopkinson pressure bar (SHPB) tests and static tensile tests [2,30]. For example, the benchmark value of the five parameters are shown in Table 1 when the workpiece material is AISI 1045. However, these values of model parameters may not conform the practical cutting process because of different cutting environment and other various factors, so it is necessary to optimize these model parameters [4]. Considering the cost and time of experiments, orthogonal experiment is adopted to select the best parameter combination.

Table 1. Material parameters in J-C model of workpiece material AISI 1045

	A (MPa)	B (MPa)	C	n	m	T_{room} (°C)	T_{melt} (°C)
Value	553.1	600.8	0.0134	0.23	1	20	1460

For the metrics, Kullback-Leibler (KL) [5,8] divergence and cosine similarity are used as the evaluation indexes of the orthogonal experiment. The KL divergence measures the difference in probability distribution of two groups of signal, and the closer the value of KL is to 0, the more similar the two groups of signal are. The formula of KL divergence is as follows [5]:

$$D_{KL}[p \,\|\, q] = \int_{-\infty}^{+\infty} p(X)\log\frac{p(X)}{q(Y)}dx \tag{2}$$

where $p(X)$ and $q(Y)$ represent the probability density of two groups of signal, respectively. The cosine similarity evaluates the similarity of two groups of signal through calculating their cosine value, and the calculation formula is as follows:

$$\cos(\theta) = \frac{X \bullet Y}{\|X\| \times \|Y\|} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{n}} X_i \times Y_i}{\sqrt{\sum_{\mathrm{i=1}}^{\mathrm{n}} (X_i)^2} \times \sqrt{\sum_{\mathrm{i=1}}^{\mathrm{n}} (Y_i)^2}} \tag{3}$$

The closer the value of $\cos(\theta)$ is to 1, the more similar the two groups of signal are. In actual engineering, it is generally considered that $\cos(\theta) > 0.6$ meets the requirements [29].

2.3. Framework of the proposed method

In this paper, a new TCM method based on numerical simulation is proposed to compensate missing samples and expand sample size. The framework of the proposed method is show in Figure 1, and the three steps of the proposed TCM are outlined in detail as follows.

Step 1: Conduct a limited number of milling TCM experiments to obtain measured cutting force signal samples, and obtain the best parameter combination of the J-C model under normal wear state of tool .

First, cutting force signal data is obtained in the milling experiments under several selected tool wear conditions by means of a three-component dynamometer. Second, the numerical model based on the J-C model is built in DEFORM, and the best parameter combination is selected by the orthogonal experiment with the comparative analysis of the simulation signal and the experimental signal under the normal tool condition, in which the criteria is minimize KL divergence satisfying the $\cos(\theta) > 0.6$.

Step2: Simulate missing sample and obtain complete wear training samples.

Table 2. Feature parameters and the calculation formulas

Domain	Feature parameter	Formula	Remarks
Time domain	Average	$T_1 = \sum_{i=1}^{n} x_i / n$	x_i (i = 1, 2, ..., n) represents the original collected signal sequence.
	Root mean square	$T_2 = \sqrt{\sum_{i=1}^{n} x_i^2 / n}$	
	Standard deviation	$T_3 = \frac{1}{n-1}\sqrt{\sum_{i=1}^{n}(x_i - T_{avg})^2}$	
	Crest factor	$T_4 = \max\{\lvert x_i \rvert\} / T_{rms}$	
	Shape factor	$T_5 = T_{rms} / T_{avg}$	
	Skewness	$T_6 = (\sum_{i=1}^{n}(x_i - T_{avg})^3) / (n \cdot T_{sd}^3)$	
	Kurtosis	$T_7 = (\sum_{i=1}^{n}(x_i - T_{avg})^4) / (n \cdot T_{sd}^4) - 3$	
	Waveform	$T_8 = n \cdot T_{rms} / (\sum_{i=1}^{n}\lvert x_i \rvert)$	
	Margin factor	$T_9 = n^2 \cdot \max\{\lvert x_i \rvert\} / (\sum_{i=1}^{n}\lvert x_i \rvert)^2$	
Frequency domain	Mean of power spectrum	$F_1 = \sum_{j=1}^{n} P_j / n$	f_i represents the frequency corresponding to x_i; P_i represents the power spectrum of f_i.
	Root mean square of power spectrum	$F_2 = \sqrt{\sum_{j=1}^{n} P_j^2 / n}$	
	Crest factor of power spectrum	$F_3 = \max\{P_i\} / \sqrt{\sum_{j=1}^{n} P_j^2 / n}$	
	Modified equivalent bandwidth	$F_4 = \sqrt{(\sum_{j=1}^{n}(f_j - \overline{f})^2 P_j) / (\sum_{j=1}^{n} P_j)}$	
	High-low ratio of power spectrum	$F_5 = (\sum_{j=n/4}^{n/2} P_j) / (\sum_{j=1}^{n/4} P_j)$	
	Stabilization ratio	$F_6 = (\sum_{j=1}^{n} f_j^2 P_j) / (\sqrt{\sum_{j=1}^{n} P_j}\sqrt{\sum_{j=1}^{n} f_j^4 P_j})$	
	Skewness of bandpower	$F_7 = (\sum_{j=1}^{n}(P_j - F_{mps})^3) / (\sum_{j=1}^{n}(P_j - F_{mps})^2)^{\frac{3}{4}}$	
	Kurtosis of bandpower	$F_8 = n(\sum_{j=1}^{n}(P_j - F_{mps})^4) / (\sum_{j=1}^{n}(P_j - F_{mps})^2)$	
Time-frequency domain	Wavelet energy coefficient	$E_j = \frac{1}{n}\sum_{k=1}^{n}(d_{j,k})^2 = \frac{1}{n}\sum_{k=1}^{n}(\int w_{j,k}(t)x(t)dt)^2$	

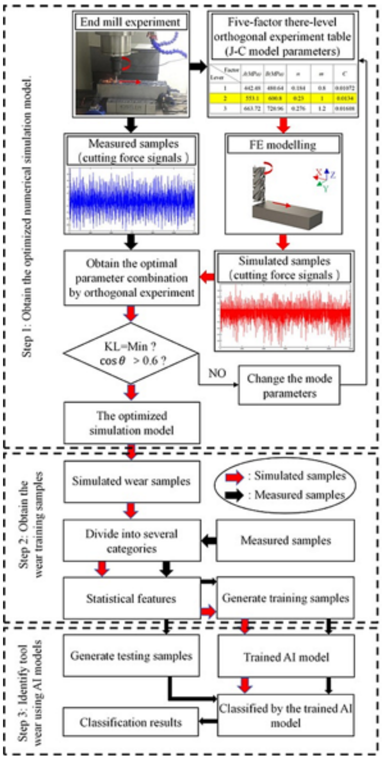


Fig. 1. Framework of the proposed TCM method

Missing tool wear categories can be defined from the experimental results, that is, these categories not occurred in experiments are missing tool wear categories. These missing tool wear categories can be simulated based on the optimal numerical model above, and the corresponding cutting force signal could be obtained. After supplementing missing samples, several feature parameters (shown in Table 2) of time, frequency, and time- frequency domains (wavelet energy coefficient) for each sample are extracted to form a feature parameter set [7, 25, 37, 48]. Here, the time-frequency domain parameter was obtained using the three-layer wavelet packet transform (WPT) with the Daubechies 2 (db2) wavelet basis function.

Step 3: Identify tool wear condition through AI classifiers.

The training set consists of simulated samples and measured samples, and inputs to train certain AI model. The trained AI model can be employed to identify unknown wear condition of tools.

3. Experimental investigations

3.1. Description of experiments

The experimental setup for the end milling TCM experiments under various operating conditions is illustrated in Figure 2. The experimental platform was built on a DMTG VDL850A vertical machining center as shown in Figure 2(a). The tools used in the experiments were uncoated three-flute tungsten steel end milling cutters (Φ 10 mm), and the workpiece material was AISI 1045 steel with dimensions of 300 mm × 100 mm × 80 mm. A three-component dynamometer (Kistler 9139AA) was mounted between the workpiece and the machine table to measure the cutting forces in the form of charges (shown in Figure 2(b)). The cutting force signal (Axial force, radial force and tangential force) was collected by a charge amplifier (Kistler 5073 A4) and a data acquisition instrument (Kistler 5697 A1) with a sampling frequency of 12 kHz (shown in Figure 2(c)). As shown in Figure 2(d), the flute wear of each cutting tool was measured after each machining stage using a GP-300C optical microscope, which represented individual milling stages. It is noteworthy that we found the influence of the length of rake face wear (KB) on the surface roughness of the workpiece after

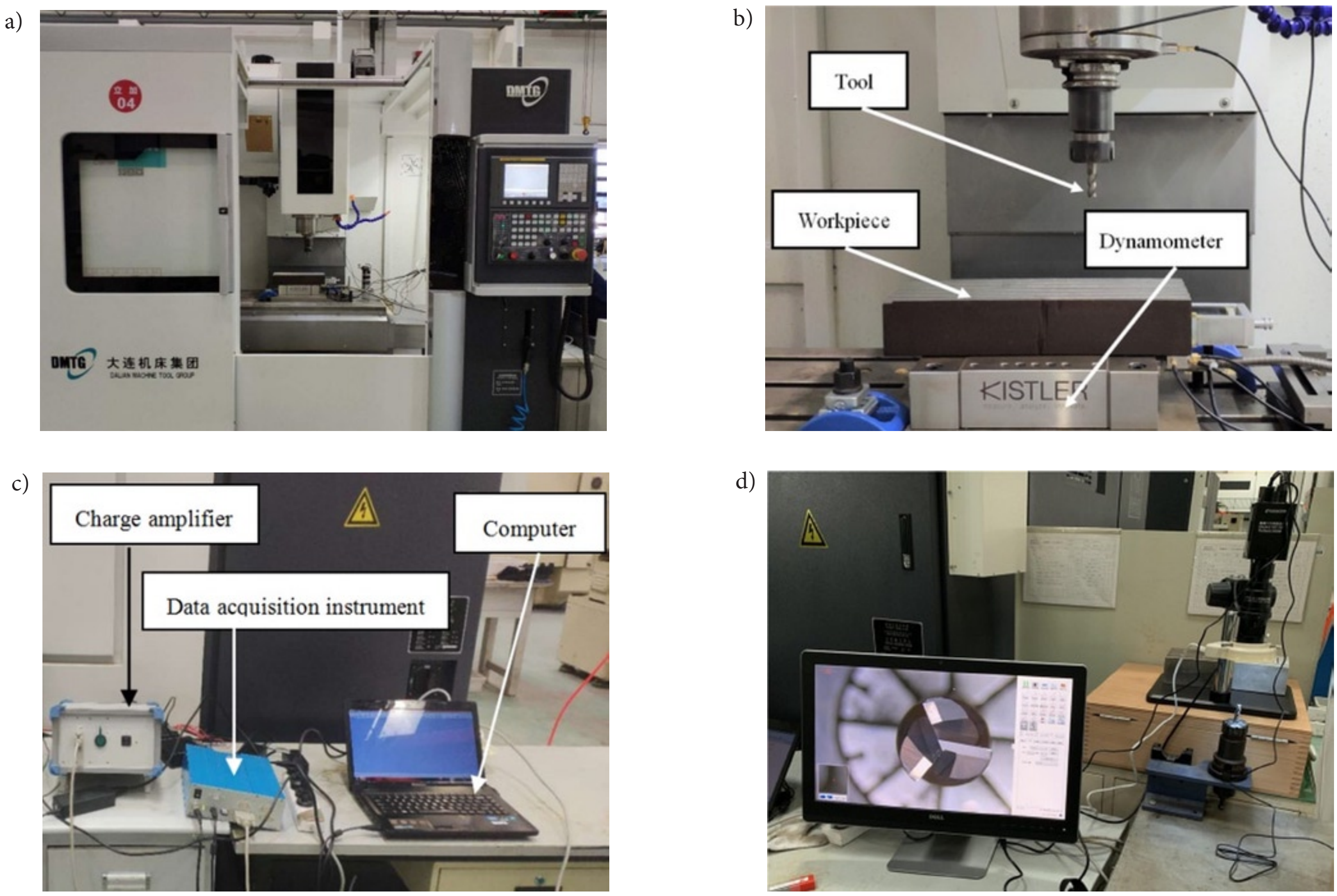


Fig. 2. Experimental end milling TCM setup [25]: a) vertical machining center, b) end milling experimental platform, c) data acquisition instrument, d) tool microscope

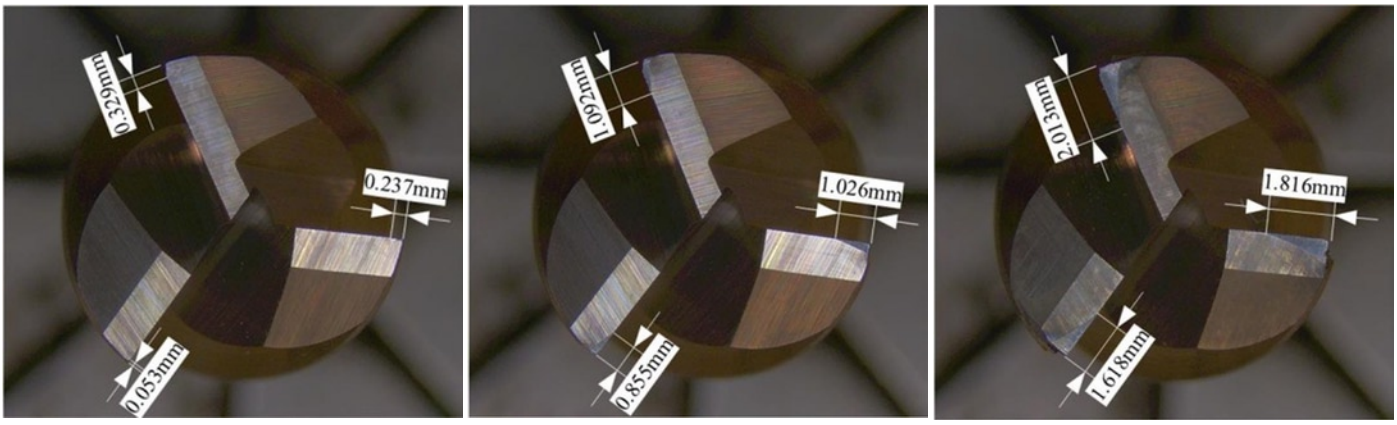


Fig. 3. Tool images indicative of different length of rake face wear (KB) values [46]: a) first milling stage, b) fifth milling stage, c) tenth milling stage

Table 3. Experimental cutting parameters

Case	Spindle Speed (rpm)	Depth of Cut (mm)	Feed Rate (mm/min)
1	2300	0.4	400
2	2300	0.6	500
3	2400	0.4	450
4	2400	0.5	500
5	2500	0.5	400
6	2500	0.6	450
7	2300	0.4	500
8	2300	0.6	400

Table 4. Tool wear classifications of the eight milling tools

Case \ Category	1-st [0, 0.3)	2-nd [0.3, 0.6)	3-rd [0.6, 0.9)	4-th [0.9, 1.2)	5-th [1.2, 1.5)	6-th [1.5, 1.8)	7-th ≥1.8	Sample number
1	—	2	2	2	—	2	2	10
2	1	1	1	3	1	3	—	10
3	—	2	1	3	1	2	1	10
4	—	2	2	2	1	3	—	10
5	1	2	2	3	2	—	—	10
6	—	2	2	2	1	1	2	10
7	2	2	3	1	2	—	—	10
8	1	1	2	1	1	4	—	10

milling was greater than that of flank wear (VB) and the depth of rake face wear (KT) [47]. Therefore, KB was employed as the tool wear criterion in the experiments, and the tool wear value after each cutting stage was defined as the maximum KB value of the three teeth. Figure 3 illustrates the progression of tool wear after finishing a single workpiece surface 1, 5, and 10 times (i.e., 1, 5, and 10 milling stages).

The experimental measurements employed eight operational conditions comprising random combinations of three operational parameters: spindle speed, depth of cut, and feed rate. The operational parameters employed in the experiments are listed in Table 3. Each case began with a new tool under the eight operational conditions and ran 10 milling stages, and the largest tool wear value obtained after completing those milling stages in all eight conditions was 2.054 mm. Therefore, the milling tool wear condition was divided into 7 categories according to tool wear intervals of 0.3 mm, and the numbers of samples observed for all conditions in all categories are listed in Table 4. It can be found in Table 4 that samples indicative of individual tool wear categories were not always available under all cutting conditions. These represent missing samples.

3.2. Numerical simulation of end milling process

First, simulation modeling was carried out according to the dimensions of workpiece and milling tool in the experiment, then the models are imported into DEFORM for processing. Second, the general pre-processing module of DEFORM was selected in the main interface, and the unit standard was set as SI. The workpiece was set as a plastic body and the material was set as AISI 1045. The tool was set as a rigid body and the material was set as tungsten carbide steel. The mechanical characteristics of these materials were imported from the rich material library in DEFORM. Then, the J-C model was selected for the workpiece material model, and the benchmark parameters of J-C model are shown in Table 1. The number of meshes for the workpiece and tool were 40,000 and 10,000, respectively. Considering the efficiency of remeshing during calculation to reduce the time of the entire milling simulation calculation, the mesh type in the model was set to a tetrahedral mesh. And the mesh size could be set to 1/3 of feed rate per spindle speed [31,39], thus according to Table 3 the mesh size could be calculated to 0.053 mm (400/2500/3 = 0.053mm). Reasonable simulation speed and accuracy was ensured by applying local refinement to the machined surface, and the refinement ratio was 0.01. After inspection, the maximum mesh size of the workpiece and the tool is less than 1/5 of the feed. Figure 4(a), (b) and (c) show the milling tool model, meshing refinement, and simulation running in DEFORM, respectively.

For boundary conditions, the bottom of the workpiece was fixed in the three directions (X, Y, and Z), the entire surface of the workpiece and tool were selected for heat exchange with the environment, the three operational parameters (spindle speed, depth of cut, and feed rate) of cutting processing were set in according to actual conditions in Table 3. The number of simulation steps was set 24000, and the sampling interval was 8.33×10^{-5} s and the sampling time was 1 s. In modeling the tool/workpiece contact, the friction coefficient between the tool and the workpiece was 0.6 [10], and the thermal conductivity was 45 $W{\cdot}m^{-1}{\cdot}C^{-1}$ [11]. Finally, after simulation, the cutting force data was exported and saved in the post-processing.

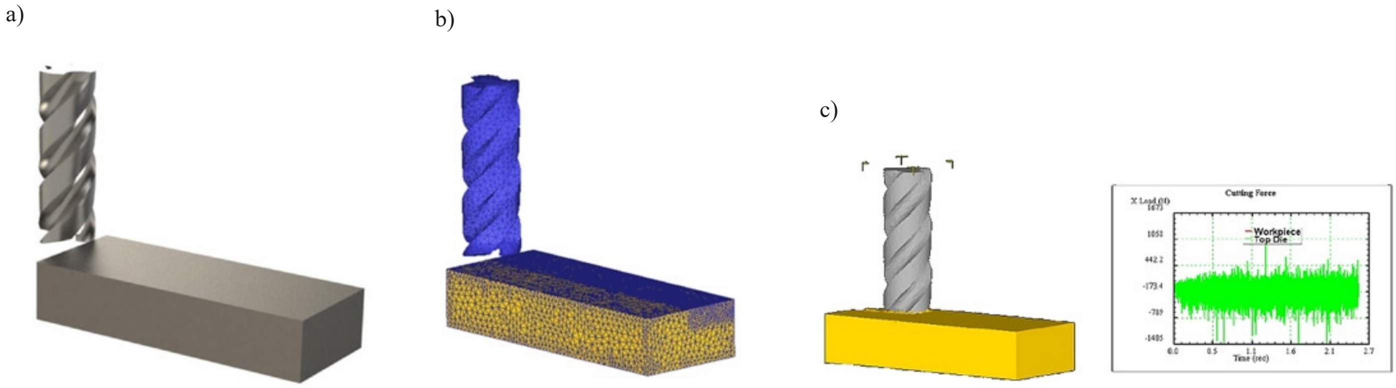


Fig. 4. Simulation of end milling process: a) milling tool model, b) meshing and refinement, c) simulation running

Table 5. Factor level table

Level	A	B	n	m	C
1	442.48	480.64	0.184	0.8	0.01072
2	553.1	600.8	0.23	1.0	0.0134
3	663.72	720.96	0.276	1.2	0.01608

Table 6. Orthogonal experiments of the J-C model parameters

No.	A	B	n	m	C	Average KL	Average Cos(θ)
1	1	1	1	1	1	2.9240	0.6991
2	1	2	2	2	2	2.8078	0.7233
3	1	3	3	3	3	3.0946	0.6942
4	2	1	1	2	2	2.6832	0.7299
5	2	2	2	3	3	2.9804	0.7236
6	2	3	3	1	1	3.0403	0.7211
7	3	1	2	1	3	2.8987	0.6954
8	3	2	3	2	1	2.8236	0.7013
9	3	3	1	3	2	3.1400	0.7136
10	1	1	3	3	2	3.1034	0.7165
11	1	2	1	1	3	3.0441	0.6957
12	1	3	2	2	1	3.0547	0.7229
13	2	1	3	1	1	2.8985	0.7366
14	2	2	1	2	2	2.9507	0.7101
15	2	3	2	3	3	2.8640	0.7099
16	3	1	2	2	3	2.9140	0.6970
17	3	2	3	3	1	2.8981	0.6996
18	3	3	1	1	2	2.9098	0.6942
Average KL of the 1-st level	3.0048	2.9036	2.9420	2.9526	2.9399	——	——
Average KL of the 2-nd level	2.9029	2.9175	2.9199	2.8723	2.9325	——	——
Average KL of the 3-rd level	2.9307	3.0172	2.9764	3.0134	2.9660	——	——
Benchmark	2	2	2	2	2	2.7516	0.7196
The optimal	2	1	2	2	2	2.6035	0.7389

3.3. Parameter optimization by orthogonal experiments

In this section,three levels of each parameter in Table 1 were set as 80%, 100% and 120% of the benchmark value (shown in Table 5), and an orthogonal table of five factors and three levels (L_{18} (5^3)) was employed to conduct the orthogonal experiments, as shown in Table 6. The operational parameters (speed, depth of cut, and feed rate) used in the orthogonal experiment were 2500 rpm, 0.6 mm, and 450 mm/min, respectively. Therefore, the experimental data of the same parameters were selected to calculate the KL divergence and cosine similarity, and the comparison data is taken one second (12,000 data points) after the milling tool completely entered the workpiece. The KL divergence and cosine similarity results of the 18 cutting tests are shown in Table 6.

The KL divergence and cosine similarity results of each orthogonal experiment case were presented in Table 6, in which the values of KL divergence and cosine similarity are the average of three directions (X, Y and Z). By main effect analysis, the best parameter combination is A(2) B(1) n(2) m(2) C(2), as shown in Table 6, the corresponding average KL divergence is 2.6035, which is smaller than the other combinations, and the average cosine similarity values (= 0.7389) greater than 0.6. The simulated and measured time series data corresponding to the best parameter combination under normal tool condition are presented in Figure 5 in the *X*, *Y*, and *Z* directions. It can be seen that the simulated signals differ slightly from the measured signals.

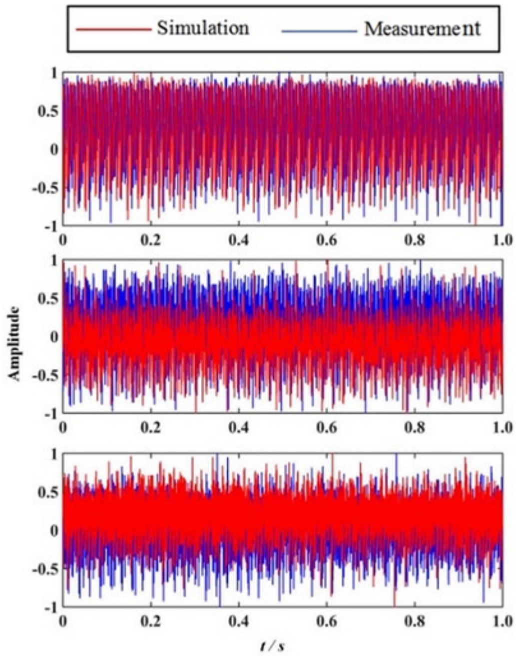


Fig. 5. The time-domain comparison between measured and simulation signals

4. Result analysis

4.1. Simulation signal verification

The validity of the simulated samples were tested by comparing 1.0 s (i.e., 12,000 sampling points) of the simulated and measured cutting force signals obtained for operational condition 6 under different tool wear categories. Figures 6-8 show the time series data and the corresponding frequency spectra of the simulated and measured signal in the *X*, *Y*, and *Z* directions under the 2-nd, 4-th and 10-th tool wear categories. It can be seen that, under these tool wear categories, the simulated signals differ slightly from the measured signals in terms of the amplitudes of the peaks in the frequency domain, while the frequency peak positions agree well.

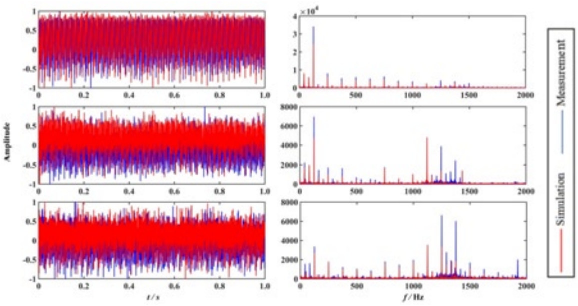


Fig. 6. Comparison of the simulated and measured cutting force signals under the 2-nd tool wear category

4.2. Sample augmentation

As shown in Figure 6, different tool wear states were simulated based on the optimal numerical simulation model according to the tool wear lengths and wear shapes obtained during the experiments, and the linear interpolation method was applied to achieve KB values less than the threshold for missing categories according to the observed tool wear value trends. Three examples of wear categories

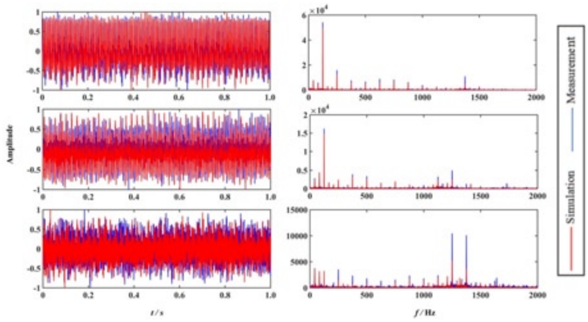


Fig. 7. Comparison of the simulated and measured cutting force signals under the 4-th tool wear category

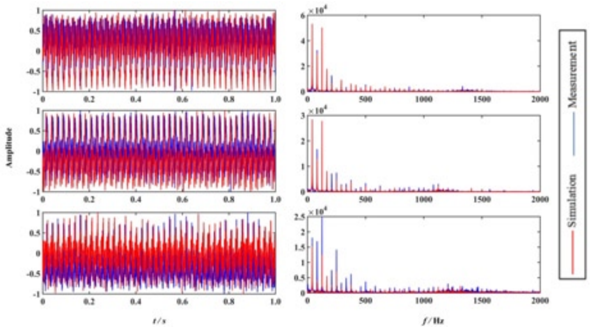


Fig. 8. Comparison of the simulated and measured cutting force signals under the 10-th tool wear category

added based on the FEM model are presented in Fig. 6 for operational condition 6.

According to Table 4, the 1-st and 7-th wear categories were generally missing under the operational conditions considered. Therefore, we consider missing samples only for the 7-th category here owing to article length limitations. Cases 2, 4, 5, 7, and 8 were employed as the training dataset because all of these are missing the 7-th wear category, and the remaining three cases 1, 3, and 6, which contain the 7-th category but not the 1-st category, were employed as the testing dataset. Then, 12,000 data points (1 s) were selected for the simulated and measured samples of each category, which are evenly divided into 20 groups. The optimal numerical simulation model model is employed to simulate the testing cases (Cases 1, 3, and 6) to increase the number of samples in the training dataset. Each simulated case contains 12 different tool KB samples involving all wear categories, and the sample sizes of the measured and simulated training sets were 900 (45×20) and 560 (28×20) not including the 1-st category, respectively. Accordingly, we employed three separate datasets to train the AI classifiers, which included the measurement dataset composed of only measured samples, the simulation dataset composed of only simulated samples, and the measurement + simulation dataset composed of measured and simulated samples, with a total of 1460 (900+560) samples.

4.3. Classification result and analysis

These feature parameters listed in Table 2 were calculated for the individual samples in the training and testing datasets, and employed as the input parameters for training and testing classifiers. Four common algorithms, SVM, RF, DT, and a generalized regression neural network (GRNN), were adopted to verify the generalized ability of the proposed method. Here, the SVM classifier selects the radial basis kernel function, and the penalty factor and kernel function radius are set to 3 and 1, respectively. The RF classifier was executed with the Randomforest-matlab open source toolbox developed by Abhishek Jaiantilal (https://github.com/ajaiantilal/ randomforest-matlab), and the number of decision trees was set to 500. The DT classifier was used the toolbox function ClassificationTree.fit in MATLAB R2016, and the parameters of 'name' and 'value' were selected as 'model' and

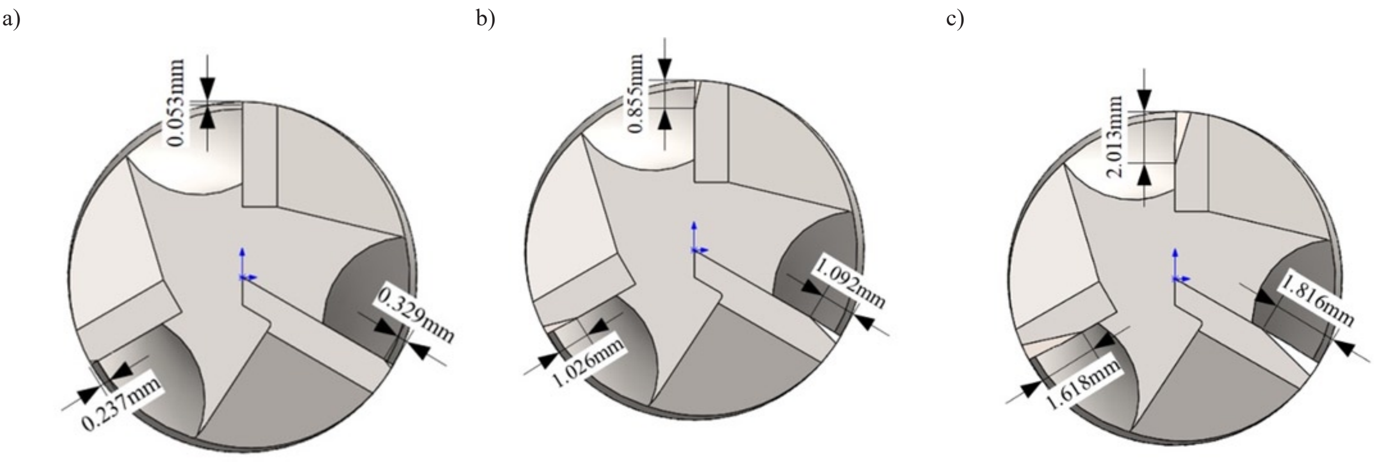


Fig. 9. Artificially added wear categories obtained from the FEM model: a) second category, b) fourth category, c) seventh category

Table 7. Classification Accuracy of four classifiers with different samples

Training set	Measure-ment	Simulation	Measurement + Simu-lation
SVM	68.67%	85.67%	91.33%
RF	73.50%	90.17%	93.83%
DT	70.00%	56.67%	90.00%
GRNN	54.67%	80.00%	83.00%
Average accuracy	66.71%	78.13%	89.54%

'graph', respectively. The value of SPREAD in the GRNN classifier was set to 0.1.

Table 7 shows the classification accuracies of four classifiers with the testing dataset. It can be found from Table 7, that the average classification accuracy obtained by the classifiers based on the simulation dataset is greater than that based on the measurement dataset by 11.42%, although the sample size of the simulation datset is less than that of the measurement dataset. There are two reasons for this result, one is the simulation dataset makes up missing categories not occurred in experiments, the other is the cutting conditions corresponding to the simulation dataset are consistent with that to the testing dataset. In addition, the average classification accuracy obtained by the classifiers based on the measurement + simulation dataset is greater than that based on the measurement dataset by 22.83%, and the classification accuracies obtained by the SVM, RF, and DT classifiers based on the measurement + simulation dataset are above 90%. Therefore, it can be considered that the proposed TCM method can improves significantly the classification accuracies of many classifiers.

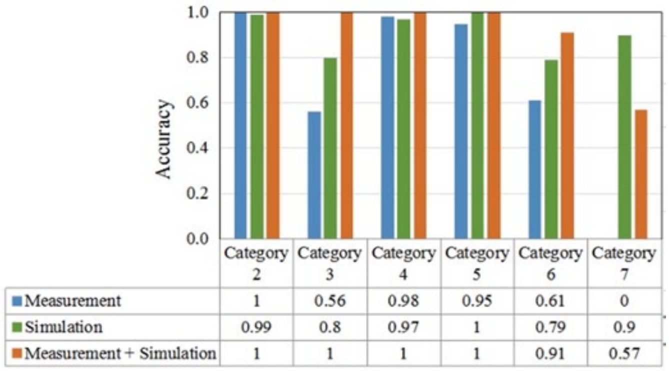


Fig. 10. Classification accuracy of each wear category using the SVM with three training datasets

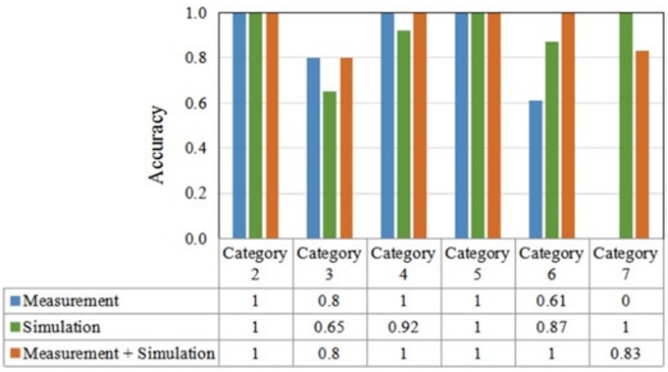


Fig. 11. Classification accuracy of each wear category using the RF with three training datasets

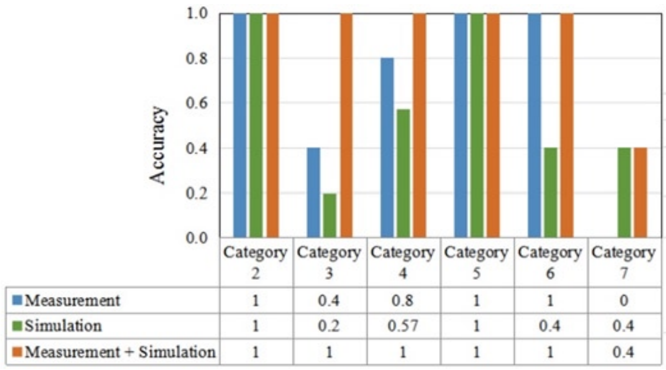


Fig. 12. Classification accuracy of each wear category using the DT with three training datasets

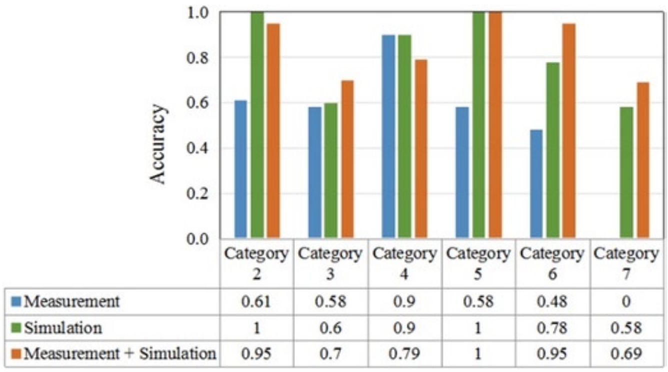


Fig. 13. Classification accuracy of each wear category using the GRNN with three training datasets.

The classification accuracy of each wear category obtained using the four classifiers trained using the three different training datasets are presented in Figures 10-13, respectively. We note from the figures that the classification accuracy of the four classifiers trained with the measurement dataset is not high for most of the wear categories. In contrast, the classification accuracy of the four classifiers trained with the simulation dataset and the measurement + simulation dataset are generally much greater (except for wear category 3 in RF and wear category 4 in GRNN, which are lower).

5. Conclusion

This paper proposed a feasible TCM method for obtaining various samples of tool wear condition by numerical simulation based on J-C model to overcome the problem of sample missing and sample insufficiency in real experiments. First, a numerical model based on Johnson-Cook model is established, and the model parameters are optimized through orthogonal experiment technology with the practical experiments, in which the Kullback- Leibler divergence and cosine similarity are used as the evaluation indexes. Second, samples under various tool wear categories are obtained by the optimized numerical model above to provide missing samples not present in the practical experiments and expand sample size. The effectiveness of the proposed method is verified by its application in end milling TCM experiments. The results indicate the classification accuracies of four classifiers (SVM, RF, DT, and GRNN) can be improved significantly by the proposed TCM method, and we believe that the proposed method has similar effects on other AI classifiers. In addition, although this study is about tool wear condition monitoring approach for end milling, the proposed method is also applicable to other machining process.

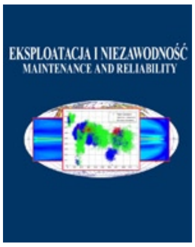
Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant 51405346, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY17E050005, and in part by the Wenzhou Key Innovation Project for Science and Technology of China under Grant ZD2019042.

References

1. Chen QP, Xie QS, Yuan QN, Huang HS, Li YT. Research on a real-time monitoring method for the wear state of a tool based on a convolutional bidirectional LSTM model. Symmetry-Basel 2019; 11(10) 1233, https://doi.org/10.3390/sym11101233.
2. Duan CZ, Yu HY, Cai YJ, Li YY. Finite Element Simulation and Experiment of Chip Formation during High Speed Cutting of Hardened Steel. Applied Mechanics and Materials 2010; 29-32: 1838-1843, https://doi.org/10.4028/www.scientific.net/AMM.29-32.1838.
3. Duan R, Lin Y, Zeng Y. Fault diagnosis for complex systems based on reliability analysis and sensors data considering epistemic uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018, 20(4): 558-566, https://doi.org/10.17531/ein.2018.4.7
4. Ducobu F, Rivière-Lorphèvre E, Filippi E. On the importance of the choice of the parameters of the Johnson-Cook constitutive model and their influence on the results of a Ti6Al4V orthogonal cutting model. International Journal of Mechanical Sciences 2017; 122: 143-155, https://doi.org/10.1016/j.ijmecsci.2017.01.004.
5. Erven T, Harremoës P. Rényi divergence and Kullback-Leibler divergence. IEEE Transactions on Information Theory 2014; 60(7): 3797-3820, https://doi.org/10.1109/TIT.2014.2320500.
6. Gao C, Xue W, Ren Y, Zhou YQ. Numerical control machine tool fault diagnosis using hybrid stationary subspace analysis and least squares support vector machine with a single sensor. Applied Sciences 2017; 7(4) 346, https://doi.org/10.3390/app7040346.
7. Gao Y, Liu XY, Xiang JW. FEM simulation- based generative adversarial networks to detect bearing faults. IEEE Transactions on Industrial Informatics 2020; 16(7): 4961-4971, https://doi.org/10.1109/TII. 2020.2968370.
8. Gkdere G, Mehmet G. New reliability score for component strength using kullback-leibler divergence. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(3): 367-372, https://doi.org/ 10.17531/ein.2016.3.7.
9. Huang ZW, Zhu JM, Lei JT, Li XR, Tian FQ. Tool wear predicting based on multi-domain feature fusion by deep convolutional neural network in milling operations. Journal of Intelligent Manufacturing 2020; 31(4): 953-966, https://doi.org/10.1007/s10845-019-01488-7.
10. Iqbal S, Mativenga PT, Sheikh MA. Characterization of machining of AISI 1045 steel over a wide range of cutting speeds. Part 1: investigation of contact phenomena. Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture 2007; 221(5): 909-916, https://doi.org/10.1243/ 09544054JEM796.
11. Iqbal S, Mativenga PT, Sheikh MA. Characterization of machining of AISI 1045 steel over a wide range of cutting speeds. Part 2: Evaluation of flow stress models and interface friction distribution schemes. Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture 2007; 221(5): 917-926, https://doi.org/10.1243/09544054JEM797.
12. Jasiulewicz-Kaczmarek M, Antosz K, Żywica P, Mazurkiewicz D, Sun B, Ren Y. Framework of machine criticality assessment with criteria interactions. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(2): 207-220, https://doi.org/10.17531/ein.2021.2.1.
13. Javidikia M, Sadeghifar M, Songmene V, Jahazi M. On the impacts of tool geometry and cutting conditions in straight turning of aluminum alloys 6061-T6: an experimentally validated numerical study. International Journal of Advanced Manufacturing Technology 2020; 106(9-10): 4547-4565, https://doi.org/ 10.1007/s00170-020-04945-3.
14. Karabacak YE, Ozmen NG, Gumusel L. Worm gear condition monitoring and fault detection from thermal images via deep learning method. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(3): 544-556, https://doi.org/10.17531/ein.2020.3.18.
15. Karandikar J, Mcleay T, Turner S, Schmitz T. Tool wear monitoring using naïve bayes classifiers. The International Journal of Advanced Manufacturing Technology 2015; 77(9-12): 1613-1626, https://doi.org/ 10.1007/s00170-014-6560-6.
16. Karwat B, Rubacha P, Stanczyk E. Simulational and experimental determination of the exploitation parameters of a screw conveyor. Eksploatacja i Niezawodnosc - Maintenance and Reliability 22(4): 741-747, https://doi.org/10.17531/ein.2020.4.18.
17. Klocke F, Lung D, Buchkremer S. Inverse identification of the constitutive equation of Inconel 718 and AISI 1045 from FE machining simulations. Procedia CIRP 2013; 8: 212-217, https://doi.org/10.1016/j. procir.2013.06.091.
18. Kong D, Chen Y, Li N, Duan C, Lu L, Chen D. Tool wear estimation in end milling of titanium alloy using npe and a novel woa-svm model. IEEE Transactions on Instrumentation and Measurement 2020; 69(7): 5219-5232, https://doi.org/10.1109/TIM.2019.2952476.
19. Kothuru A, Nooka SP, Liu R. Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling. The International Journal of Advanced Manufacturing Technology 2018; 95(9-12): 3797-3808, https://doi.org/10.1007/s00170-017-1460-1.
20. Kozłowski E, Mazurkiewicz D, Zabinski T, Prucnal S, Sep J. Assessment model of cutting tool condition for real-time supervision system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(4): 679-685, https://doi.org/10.17531/ein.2019.4.18.
21. Kozłowski E, Mazurkiewicz D, Żabiński T, Prucnal S, Sęp J. Machining sensor data management for operation-level predictive model.

Expert Systems with Applications 2020; 159: 1-22, https://doi.org/ 10.1016/j.eswa.2020.113600.
22. Kumar A, Kumar R. Adaptive artificial intelligence for automatic identification of defect in the angular contact bearing. Neural Computing and Applications 2018; 29: 277-287, https://doi.org/10.1007/s00521-017 -3123-4.
23. Kumar A, Kumar R. Least Square Fitting for Adaptive Wavelet Generation and Automatic Prediction of Defect Size in the Bearing Using Levenberg-Marquardt Backpropagation. Journal of Nondestructive Evaluation 2017; 36, 7, https://doi.org/10.1007/s10921-016-0385-1.
24. Küppers F, Albers J, Haselhoff A. Random Forest on an Embedded Device for Real-time Machine State Classification. 2019 European Signal Processing Conference (EUSIPCO), A Coruna, Spain, https://doi.org/10.23919/EUSIPCO.2019.8902993.
25. Lei Z, Zhou YQ, Sun BT, Sun WF. An intrinsic timescale decomposition-based kernel extreme learning machine method to detect tool wear conditions in the milling process. The International Journal of Advanced Manufacturing Technology 2020; 106(3-4): 1203-1212, https://doi.org/10.1007/s00170-019- 04689-9.
26. Li GF, Wang YB, He JL, Hao QB, Yang HJ, Wei JF. Tool wear state recognition based on gradient boosting decision tree and hybrid classification RBM. The International Journal of Advanced Manufacturing Technology 2020; 110: 511-522, https://doi.org/10.1007/s00170-020-05890-x.
27. Liu D, Wang S, Tomovic M. Degradation modeling method for rotary lip seal based on failure mechanism analysis and stochastic process. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(3): 381-390, https://doi.org/10.17531/ein.2020.3.1
28. Liu H, Liu ZY, Jia WQ, Lin XK, Zhang S. A novel transformer-based neural network model for tool wear estimation. Measurement Science and Technology 2020; 31(6) 065106, https://doi.org/10.1088/1361- 6501/ab7282.
29. Liu XY, Huang HZ, Xiang JW. A personalized diagnosis method to detect faults in gears using numerical simulation and extreme learning machine. Knowledge-Based Systems 2020; 105653, https://doi.org/10.1016/j.knosys.2020.105653.
30. Liu ZQ, Wu JH, Shi ZY, Zhao PF. State-of-the-art of Constitutive Equations in Metal Cutting Operations. Tool Engineering 2008; 42(3): 3-10 (in Chinese).
31. Padma RB, Kumara SM. Finite Element Simulation of a Friction Drilling process using Deform-3D. International Journal of Engineering Research and Applications 2012; 2(6): 716-721.
32. Pawełczyk M, Fulara S, Sepe M, Luca AD, Badora M. Industrial gas turbine operating parameters monitoring and data-driven prediction. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(3):391-399. https://doi.org/10.17531/ein.2020.3.2
33. Serin GB, Sener AM, Ozbayoglu HO. Review of tool condition monitoring in machining and opportunities for deep learning. International Journal of Advanced Manufacturing Technology 2020; 109(2): 953-974. https://doi.org/10.1007/s00170-020-05449-w.
34. Shankar S, Mohanraj T, Rajasekar R. Prediction of cutting tool wear during milling process using artificial intelligence techniques. International Journal of Computer Integrated Manufacturing 2019; 32(2): 174-182, https://doi.org/10.1080/0951192X.2018.1550681.
35. Shao F, Liu Z, Wan Y, Shi Z. Finite element simulation of machining of Ti-6Al-4V alloy with thermodynamical constitutive equation. The International Journal of Advanced Manufacturing Technology 2010; 49: 431-439, https://doi.org/10.1007/s00170-009-2423-y.
36. Shrot A, Bäker M. Determination of Johnson-Cook parameters from machining simulations. Computational Materials Science 2012; 52 (1): 298-304, https://doi.org/10.1016/j.commatsci.2011.07.035.
37. Su C, Chen H, Wen Z. Prediction of remaining useful life for lithium-ion battery with multiple health indicators. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(1): 176-183, 10.17531/ein.2021.1.18.
38. Sun HB, Zhang JD, Mo R, Zhang XZ. In-process tool condition forecasting based on a deep learning method. Robotics and Computer-Integrated Manufacturing 2020; 64: 101924, https://doi.org/10.1016/j.rcim. 2019.101924.
39. Tamizharasan T, Kumar S. Optimization of cutting insert geometry using DEFORM-3 D: numerical simulation and experimental validation. International Journal of Simulation Modelling 2012; 11(2): 65-76, https://doi.org/10.2507/IJSIMM11(2)1.200.
40. Wang Y, Su HH, Dai JB, Yang SB. A novel finite element method for the wear analysis of cemented carbide tool during high speed cutting Ti6Al4V process. The International Journal of Advanced Manufacturing Technology 2019; 103, 2795-2807, https://doi.org/10.1007/s00170-019-03776-1.
41. Wu DZ, Jennings C, Terpenny J, Gao RX, Kumara S. A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. Journal of Manufacturing Science and Engineering-Transactions of the ASME 2017; 139(7): 071018, https://doi.org/10.1115/ 1.4036350.
42. Xiang JW, Zhong YT. A novel personalized diagnosis methodology using numerical simulation and an intelligent method to detect faults in a shaft. Applied Sciences 2016; 6(12) 414, https://doi.org/10.3390/ app6120414.
43. Xu Z, Guo D, Wang J, Ge D. A numerical simulation method for a repairable dynamic fault tree. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(1):34-41, https://doi.org/10.17531/ein.2021.1.4.
44. Zhou J, Pang C, Zhong Z, Lewis F. Tool wear monitoring using acoustic emissions by dominant- feature identification. IEEE Transactions on Instrumentation and Measurement 2011; 60(2): 547-559, https://doi.org/10.1109/TIM.2010.2050974.
45. Zhou YQ, Sun BT, Sun WF. A tool condition monitoring method based on two-layer angle kernel extreme learning machine and binary differential evolution for milling. Measurement 2020; 16(15): 186-196, https://doi.org/10.1016/j.measurement.2020.108186.
46. Zhou YQ, Sun BT, Sun WF, Lei Z. Tool Wear Condition Monitoring Based on a Two-layer Angle Kernel Extreme Learning Machine using Sound Sensor for Milling Process. Journal of Intelligent Manufacturing 2020; 9, https://doi.org/ 10.1007/s10845-020-01663-1.
47. Zhou YQ, Sun WF. Tool Wear Condition Monitoring in Milling Process Based on Current Sensors. IEEE Access 2020; 8: 95491-95502, https://doi.org/10.1109/ACCESS.2020.2995586.
48. Zhou YQ, Xue W. A multisensor fusion method for tool condition monitoring in milling. Sensors 2018; 18(11): 3866, https://doi.org/10.3390/s18113866.
49. Zhou YQ, Xue W. Review of tool condition monitoring methods in milling processes. The International Journal of Advanced Manufacturing Technology 2018; 96(3-4): 2509-2523, https://doi.org/10.1007/s00170- 018-1768-5.
50. Zhu KP, Liu T. Online tool wear monitoring via hidden semi-Markov model with dependent durations. IEEE Transactions on Industrial Informatics 2018; 14(1): 69-78, https://doi.org/10.1109/TII.2017.2723943.



Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Mazurkow A, Witkowski W, Kalina A, Wierzba B, Oleksy M. The effect of oil feeding type and oil grade on the oil film bearing capacity. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 381–386, http://doi.org/10.17531/ein.2021.2.18.

The effect of oil feeding type and oil grade on the oil film bearing capacity

Indexed by: Web of Science Group

Aleksander Mazurkow[a], Waldemar Witkowski[a,*], Adam Kalina[a], Bartłomiej Wierzba[a], Mariusz Oleksy[a]

[a] Rzeszów University of Technology, al. Powstańców Warszawy 8, Rzeszów, Poland

Highlights

- Oil hydrodynamic bearings can be fed from a lubrication pocket or from the face side.
- The analysis for two oil fed method and two oils were made.
- The method of oil fed and type of oil depends the bearing static characteristics.
- This study could help to chose the oil and fed method for plain bearings.

Abstract

Two types of hydrodynamically lubricated plain journal bearings were subject to examination differing in the method used to feed them with oil. The first type was fed from a lubrication pocket and the second from the bearing face side. Mathematical models were developed with two-way oil flow allowing to determine the oil film bearing capacity, the maximum pressure, the maximum temperature, and the film oil minimum height for given position of journal relative to solid bush. Static characteristics were developed used in the further course of the study to compare operating parameters of the considered types of bearings. Another issue considered in the paper is the effect of oil VG grade on bearing performance with conditions of oil feeding taken into account and results of the research presented.

Keywords

hydrodynamic bearing, oil feeding method, static characteristics, viscosity, VG oil grade, relavity eccentrity, THD.

List of basic symbols

B – bush width (m); $C_R = 0.5(D - D_J)$ – radial clearance (m); $D = 2R$ – diameter (m); e – eccentricity (m); h – oil film height (m); F – load (N); h_{min} – minimum oil clearance height (m); n_J – journal rotational speed (rpm); p – pressure in oil film (N/m^2); T – temperature (°C); $x = \varphi \cdot R$ – Cartesian system coordinate (m); y – Cartesian system coordinate (m); z – Cartesian system coordinate (m); β – journal center (O_J) and bush center (O) attitude line angle; $\varepsilon = e/C_R$ – relative eccentricity; ω_J – Journal rotational speed (rad/s). Indexes: B – solid bush; J – journal, ZC – fresh oil feeding from the bearing face side; ZK – feeding with fresh oil from a lubrication pocket.

1. Introduction

Plain bearings are commonly used in various types of machine solutions or in transmission systems. Increasingly higher requirements for bearing systems, such as high durability, operation at high rotational speeds, increasing greater thermal loads, increasing load capacity, and lowering the vibration level, require an in-depth analysis of their properties. Currently, research are conducted both of a theoretical basis - mathematical models simulating real working conditions, and experimental ones [13, 17].

One of the example are results of tests of dynamic properties of the rotor bearing system presented in [4, 6, 16, 26]. Computational simulations showed that for the new bearing concept load capacity, temperature in the oil film and fluid-induced instability conditions are dependent on the rotational speed directions. For positive rotational speed of the bearing (shaft and bearing surfaces rotate in the same directions) the average velocity of the oil film (thus, the load capacity) was increasing and viscous shear of the film (reducing oil temperature) was decreasing. For opposite directions of the bearing rotational speed the average velocity of the oil film was decreasing (avoiding fluid-induced instability) [15].

The instability of hydrodynamic bearings can be diagnosed by using Teager-Kaiser energy operator. The experimental tests were conducted for two cases of rotor unbalance: G6.3 in accordance in ISO 1940-1 [28] standard and twice greater as allowed in the ISO 1940-1 standard. In the results analyses the energy operator, measured rotor displacement and acceleration of bearing were presented in [3]. By using new analytical method for calculating the nonlinear floating ring bearings oil film the unbalance effect influencing the rotor response was presented in [20]. The engine excitation effects shown that the rotor response has a distinct difference at lower and higher speeds as well.

The angular misalignment effects on bearing performance is a very important issue. The manufacturing tolerances, installation error and elastic deflection of the rotors are the misalignment potential causes. The numerical investigation of pad tilt motion and the spherical pivot of tilt, pitch and yaw motions was presented in [10, 19]. The increasing misalignment led to the stiffened bearing and decreased minimum film thickness (increased lubricant peak temperature) [19]. The effect of the axial movement of journal on the maximum film pressure, load

(*) Corresponding author.

E-mail addresses: A. Mazurkow - almaz@prz.edu.pl, W. Witkowski - wwitkowski@prz.edu.pl, A. Kalina - akalina@prz.edu.pl, B. Wierzba - bwierzba@prz.edu.pl, M. Oleksy - molek@prz.edu.pl

capacity and over-turning moment and friction power loss is relatively weak for the small eccentricity and for the smaller the rotational speed is greater [10].

The problems associated with environmental protection can be solved with the lubrication by water. The results of research on new design solutions for this type of bearings are presented in [7]. As a result of the research, it was noticed that the turbulent flow in relation to the laminar flow increases the load bearing capacity. On the other hand, an increase of the water temperature in the bearing reduces the load capacity [7]. The influences of misalignment on the lubrication performances and lubrication regimes transition of water lubricated bearing was presented in [23]. With the increase of the misaligned angle the maximum pressure and shear stress increased, the minimum film thickness decreased and the eight dynamic coefficients increased. The micro interface lubrication regime influence on the streamline, pressure, eddy viscosity and kinetic energy distribution in the micro cavities were discussed in [24].

The transverse self-aligning hydrodynamic bearings operating in the turbine drive systems at high speeds have good hydrodynamic working stability. The influence of the oil film pressure and temperature distributions on the pads deformation was presented. Tests were conducted for static equilibrium position of the journal [5, 18]. The bearing load capacity related to each pressure distribution can be calculated by researching the dimensionless lubricant film thickness in the circumference direction. The lubricant film thickness reflects directly the bearing topology structure can be expressed by harmonic functions [25].

The results of the research on increasing the bearing capacity, reducing the coefficient of friction and wear with the use of nanofluids are presented in [1, 16]. The use of TiO_2 nanoparticles in a nanofluid for different oils (DTE 26, DTE 25, DTE 24) and different rotational speeds was investigated. Increasing the rotational speed from 500 rpm to 1500 rpm caused that the dissipation power and temperature increased around 600% and 800%, respectively [1]. The use of tungsten disulfide nanoparticles (IF-WS_2 NPs) in nanofluid increased the bearing load carrying capacity about 18% [16]. Increased load carrying capacity, significantly reduced peak pressures, more oven oil film pressure distribution and thicker oil film in the loaded zone compared to a white metal bearing can be obtained also by using PTFE layer as a bearing liner [9].

The adiabatic or diathermic theoretical models, which taking into account the influence of temperature on the oil viscosity, as shown by the research results, significantly make the temperature distribution in the oil film more accuracy to the results of experimental research [8, 14]. Concerning the ability to predict friction power losses in journal bearings, the research results indicate that the considerably simpler elastohydrodynamic approach appears to be sufficient to reliably and accurately predict these losses for full film lubrication and to investigate the occurrence of metal-metal contact [2, 21].

In many mechanical application the floating ring bearings are used instead the plain bearings [12]. But theoretical models are much more complex for that bearings. To analyze mechanical and thermal performances the thermohydrodynamic model can be used [11]. The floating ring bearings system is inherently nonlinear. If it is lightly loaded or operated at high speeds, it is prone to the fluid-induced instability. Several approaches for the linearization of the forces acting in floating ring bearings were proposed and analyzed in [6].

In industrial practice, the radial plain bearings can be feed with fresh oil from a lubrication pocket located in the non-working part of the oil film or from the face side of the bearing. In the literature and in standards [27, 13], methods for calculating the bearings operating parameters feed with fresh oil from a lubrication pocket are presented. But there are no such methods for the oil feed from the face side of the bearing. The experimental results of thermal phenomena accompanying operation of a water-lubricated stern tube bearing with axial grooves (lubricant feed from face side) were discussed in [22]. It should be noted here, that face side oil feed is a common method used for example in the bearing of crankshafts, engine timing gear, and turbochargers.

The basis of mathematical models of this type of bearings are the equations of the pressure distribution in the oil and temperature distribution in the lubricating gaps, and the equation of the oil clearance geometry. The above equations are supplemented with the equation of the mathematical model of the oil lubricating the bearing. The pressure and temperature distribution equations are differential equations that are solved for boundary conditions reflecting the actual operating conditions of the bearing [1, 14]. The boundary conditions are related to the oil feed form.

The conducted analysis of the state of knowledge has shown that the issues concerning, among others: operation and construction of slide bearings are the subject of many scientific studies. However, it should be noted that in the available literature, no studies have been found concerning plain bearings feed with oil from face side. Moreover, the available literature does not present any mathematical models to describe the lubrication of bearings with such a supply. Therefore, it was considered justified to build a mathematical model for these bearings to describe the properties of the oil film, which will be based on models intended for bearings supplied with oil from the lubricating pocket. The mathematical model was the hydrodynamic lubrication model, taking into account the influence of temperature on the oil viscosity. The static characteristic of plain bearings were developed for a better presentation of the bearing working condition. These characteristic allow for the given journal position relative to the solid bushing (ε) and type of the oil $\eta(T)$ to determine the maximum pressure ($p_{\max}$), the maximum temperature ($T_{\max}$), and the oil clearance minimum height ($h_{\min}$).

Another issue, that has been considered in this manuscript, and which significantly affects on the bearing operation conditions, was the type of the oil oil viscosity used to feed the bearing. The ISO oil viscosity classification according to ISO 3448:1992 [29] was used for the tests. The obtained results allow to determine, if the oil used for plain bearing lubrication, was chosen appropriate. The maximum oil film temperature as a criterion for proper operation condition was used. The test results were presented in the form of graphs. Other parameters influencing the lubricating properties will be taken into account in further research.

2. Materials and models

2.1. Equations of mathematical model for a bearing fed from lubrication pocket

The structure, geometry, and oil flows in a bearing fed with oil from a lubrication pocket is presented in Figure 1. For the purpose of involved calculations, the method recommended in the standard [27] was used.

Similarly, the structure, geometry, and oil flows in a bearing fed from the face side are shown in Figure 2. The calculation model employed for this oil supply type was verified by means of experimental studies [13, 14].

2.2. The mathematical model constitutes a system of equations describing:

- **pressure distribution in oil clearance:**

$$\frac{4}{D^2}\frac{\partial}{\partial\phi}\left(h^3\cdot\frac{\partial p}{\partial\phi}\right)+\frac{\partial}{\partial z}\left(h^3\cdot\frac{\partial p}{\partial z}\right)=6\cdot\eta\cdot\omega_J\,\frac{\partial h}{\partial\phi} \quad (1)$$

Equation (1) was obtained after transformation of the equation of momentum conservation for oil particles and the equation of flow continuity [27].

- **oil clearance shape:**

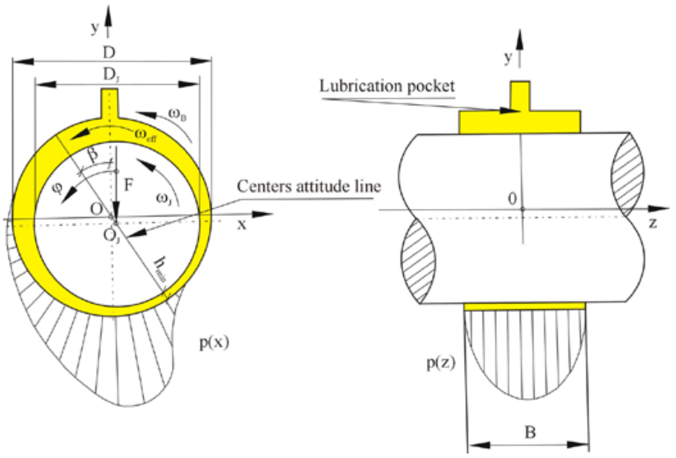


Fig. 1. The geometry, oil flow directions, and oil pressure distribution in a plain journal bearing fed with fresh oil from lubrication pocket

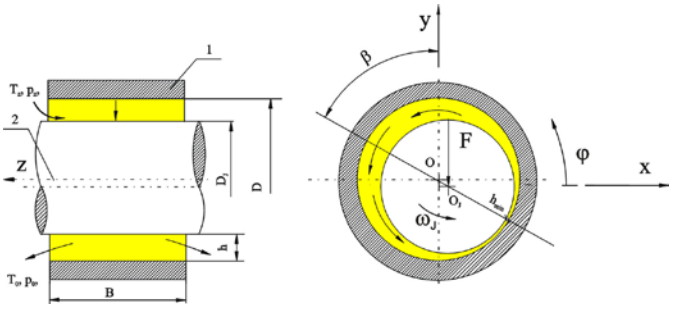


Fig. 2. The geometry, oil flow directions in case of feeding with fresh oil from the bearing face side

$$h = 0.5 \cdot D \cdot \psi_R \left[1 + \varepsilon \cdot cos(\varphi - \beta)\right] \tag{2}$$

where: $\psi_R = \dfrac{C_R}{D}$

Equation (2) was obtained by assuming the reference system as in Figure 1 or Figure 2 [27].

- **emperature distribution in oil clearance in case when the heat from bearing is carried away by flowing oil:**

$$\rho \cdot c_p \cdot \left[\upsilon_x^* \cdot \frac{\partial T}{\partial x} + \upsilon_z^* \cdot \frac{\partial T}{\partial z}\right] = \eta \cdot \left[\upsilon_x^{**} + \upsilon_z^{**}\right] \tag{3}$$

where:

$$\upsilon_x^* = \int_0^h \upsilon_x dy, \upsilon_z^* = \int_0^h \upsilon_z dy, \upsilon_x^{**} = \int_0^h \left(\frac{\partial \upsilon_x}{\partial y}\right)^2 dy, \upsilon_z^{**} = \int_0^h \left(\frac{\partial \upsilon_z}{\partial y}\right)^2 dy.$$

The equation describing the temperature distribution was derived from the energy balance equation.

- **properties of oil lubricating the bearing:**

$$\eta(T) = \eta_0 \cdot e^{a_\eta \cdot (T - T_0) + b_\eta \cdot (T - T_0)^2}, \ \rho(T) = const, \ c_p(T) = const \tag{4}$$

For the sake of the present considerations it is assumed that the oil is a Newtonian fluid [27].

In case of the bearing fed from lubrication pocket, the system of mutually adjoint equations (1–4) was solved for boundary conditions applicable to the pressure field and the temperature field. The conditions are represented in Figure 3.

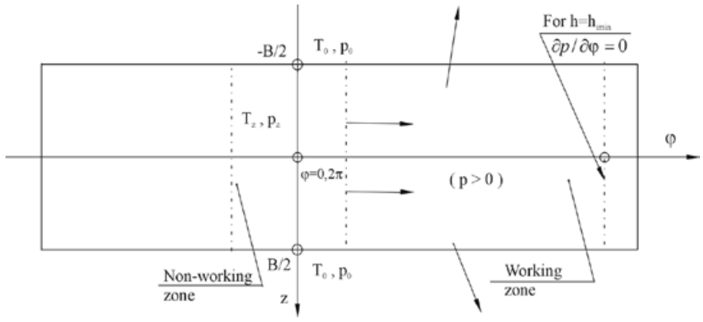


Fig. 3. Boundary conditions for pressure and temperature field in the model of bearing fed with oil from a lubrication pocket. Lines ended with arrows indicate directions of oil flow in the bearing

In case of the bearing fed with oil from its face side, the system of Equations (1–4) was solved for boundary conditions applicable to the pressure field. The conditions are presented in Figure 4.

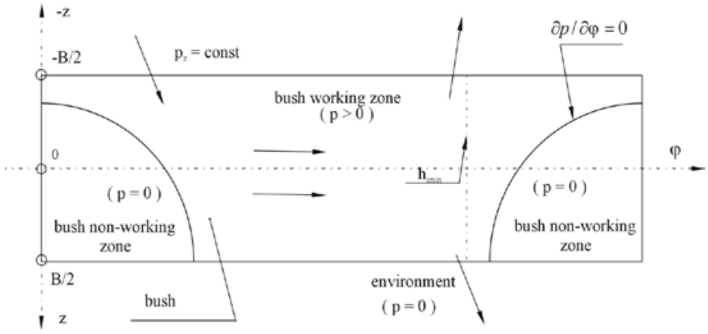


Fig. 4. The pressure field boundary conditions for a bearing fed with oil from the face side. Lines ended with arrows indicate directions of oil flow in the bearing

On the other hand, boundary conditions for the temperature field are depicted in Figure 5, where:

$$\beta_x = \frac{\eta\left(\upsilon_{xi}^{**} + \upsilon_z^{**}\right) \cdot \upsilon_{xi}^*}{\rho \cdot c_p((\upsilon_z^*)^2 + \left(\upsilon_{xi}^*\right)^2)} \quad \beta_z = \frac{\eta\left(\upsilon_{xi}^{**} + \upsilon_z^{**}\right) \cdot \upsilon_z^*}{\rho \cdot c_p((\upsilon_z^*)^2 + \left(\upsilon_{xi}^*\right)^2)} \tag{5}$$

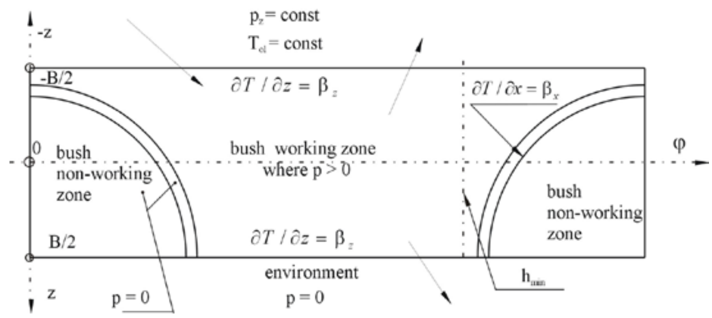


Fig. 5. The temperature field boundary conditions for a model of bearing fed with oil from the face side

The result of solving the problem of thermo-hydrodynamic equilibrium of the journal relative to the solid bush are the following quantities: $p(\phi,z), T(\phi,z), h(\phi,z), F = F_L$

3. Results

3.1. A comparative study on operating parameters of bearings fed with oil from a lubrication pocket and from the face side

For the present study, two oils have been selected with properties summarized in Table 1. The calculations were carried out for a bearing fed with oil either from lubrication pocket or from the bearing face side. As a preset quantity, position of the journal relative to the solid bush (ε) was assumed. Results of the research in the form of static characteristics as functions depending on the oil type are presented in graphical form in Figure 6.

Table 1. Preset quantities

	Preset parameters	
1.	Journal nominal diameter	D_J = 131.925 mm
2.	Solid bush nominal diameter	D = 132.109 mm
3.	Journal-floating bush relative clearance	ψ = 1.39‰
4.	Relative width	B/D = 0.5
5.	Relative eccentricity	$\varepsilon = \langle 0.2–0.85 \rangle$
6.	Journal rotational speed	ω_J= 500 s^{-1}, n_1= 4774.65 rpm
7.	Oil viscosity	$\eta_0^{(1)}$ = 0.1084 Pa·s, $a_\eta^{(1)} = -55291\cdot10^{-6}$, $b_\eta^{(1)} = 239\cdot10^{-6}$ $\eta_0^{(2)}$ = 0.5264 Pa·s, $a_\eta^{(2)} = -75000\cdot10^{-6}$, $b_\eta^{(2)} = 349\cdot10^{-6}$
8.	Oil density	ρ_0 = 900 kg/m^3
9.	Oil specific heat	c_{p0} = 2000 J/kg·°C
10.	Bearing feeding oil and environment temperature	T_z = 50°C, T_0 = 20°C
11.	Bearing feeding oil pressure	p_z = 0.1 MPa

Table 2. Bearing operating parameters

Oil feeding method	p_{allow}=15 [MPa]	
	η_0=0.1084 [Pa·s]	η_0=0.5264 [Pa·s]
Oil feeding from the bearing face side	ε^p_{allow}=0.81 T_{max}=111[°C]	ε^p_{allow}=0.75 T_{max}=124[°C]
Oil feeding from a lubrication pocket	ε^p_{allow}=0.75 T_{max}=117[°C]	ε^p_{allow}=0.71 T_{max}=130[°C]

The relative eccentricity (Figure 6, Tables 2 and 3) determining position of the journal relative to the solid bush (ε) has an effect on fulfillment of criteria of current operation, namely: $p(\varphi,z) \le p_{allow}$, $T(\varphi,z) \le T_{allow}$, $h_{min} \le h_{allow}$.

With increasing value of ε (Figure 6), value of the maximum pressure in oil film (p_{max}) increases accordingly. Adopting the allowable value of surface pressures for the bush material p_{allow} =15 MPa, values allowable for the relative eccentricity ε^p_{allow} for the lubrication pocket oil feeding method are lower than those in the case of oil being fed from the bearing face side (Figure 6, Table 2).

It follows from analysis of the course of the maximum temperature function (Figure 6) that with increasing value of ε, the maximum temperature initially decreases, and for $\varepsilon \ge 0.5$ it starts to increase. The function $T_{max} = T_{max}(\varepsilon)$ for both of the two feeding methods reaches its minimum for the value $\varepsilon^{Tmax}_{min} \approx 0.45$. Adopting T_{allow} = 95°C as the allowable oil temperature, allowable values of the relative eccentricity ε^T_{allow} , the corresponding oil film bearing capacity values ($F_{Lallow} = F_{allow}$), and maximum pressures with the feeding method taken into account the values which are presented in Table 3. For an oil with η_0 = 0.5264 Pa·s, values of the maximum temperature ($T_{max} > T_{allow}$ = 95°C) are exceeded in the whole examined range.

With increasing value of ε, the minimum oil clearance height decreases. For the discussed calculation example, the condition of fluid friction is met in the whole range of considered relative eccentricity values ($h_{min} \le h_{allow}$ =15 μm).

3.2. The effect of oil VG grade and oil feeding method on operating conditions of the bearing

In order to determine the effect of oil grade, oils of the following viscosity grades were examined: VG32, VG46, VG68, VG100, VG150. Results of tests allowing to establish the effect of oil grade and the oil feeding method are presented in graphical form as plots of the function $T_{max} = T_{max}(VG, \varepsilon, F_L$, oil feeding method). For the assumed preset values, the function takes the form shown in Figure 7.

Table 3. Operating parameters of the bearing

Oil feeding method	T_{allow} = 95°C	
	η_0 = 0.1084 Pa·s	η_0 = 0.5264 Pa·s
Oil feeding from the bearing face side	ε^T_{allow} = 0.71 F_{allow} = 23.5 kN p_{max} = 8.0 MPa	$T_{max} > T_{allow}$
Oil feeding from a lubrication pocket	ε^T_{allow} = 0.64 F_{allow} = 19.0 kN p_{max} = 7.0 MPa	$T_{max} > T_{allow}$

For the purpose of the study, two different values of the relative eccentricity ε were adopted, namely 0.45 and 0.7.

By analyzing the course of the function $T_{max} = T_{max}(\varepsilon, T_{allow}, F_L$, VG) and taking into account the fresh oil feeding method (Figure 7), a significant effect of oil VG grade and feeding method can be noted on bearing operation parameters such as the oil film bearing capacity or the maximum oil temperature.

4. Summary

Characteristics were developed allowing to determine conditions for correct operation of a bearing with oil VG grade taken into account. The effect of the oil viscosity, the oil clearance geometry, and the oil film pressure and temperature was demonstrated for two oil feeding methods.

By analyzing the research results presented in Table 3 it was found that for the same geometrical parameters, the face-fed bearings have higher bearing capacity. For the discussed structural design solution, the relative bearing capacity increase was found to be dF_{Lallow} = 18%.

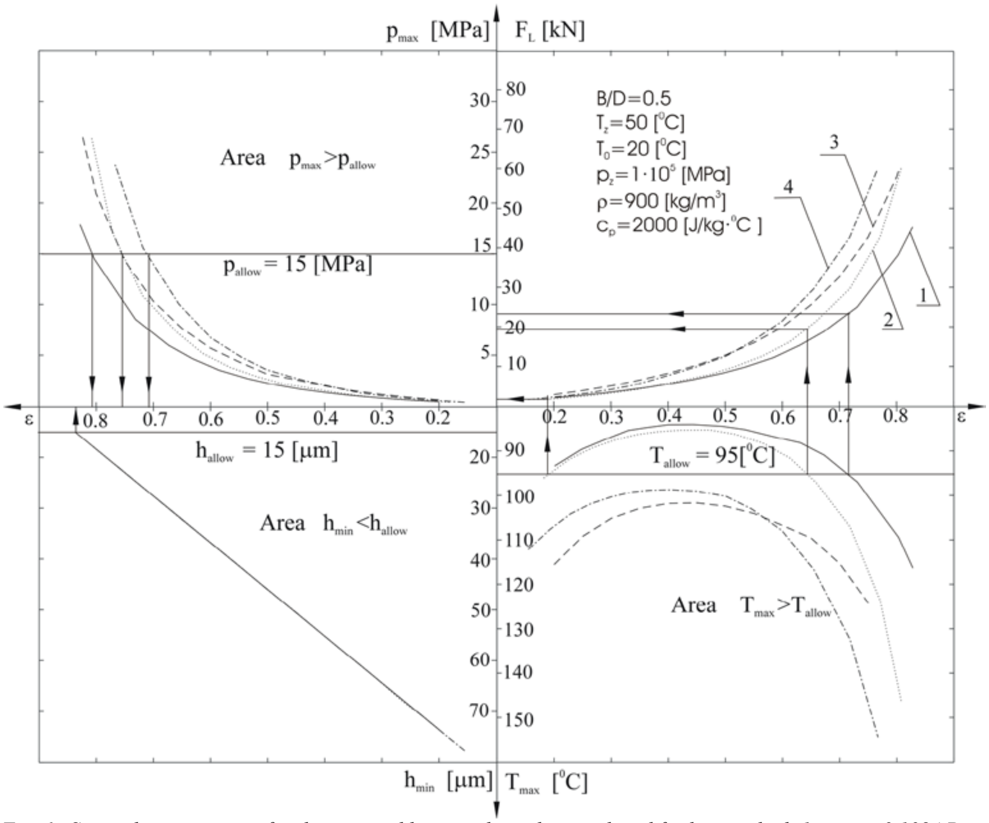


Fig. 6. Static characteristics of a plain journal bearing depending on the oil feeding method: 1 — η_0 = 0.1084 Pa·s, feeding from the bearing face side; 2 — η_0 = 0.1084 Pa·s, feeding from a lubrication pocket; 3 — η_0 = 0.5264 Pa·s, feeding from the bearing face side; 4 — η_0 = 0.5264 Pa·s, feeding from lubrication pocket

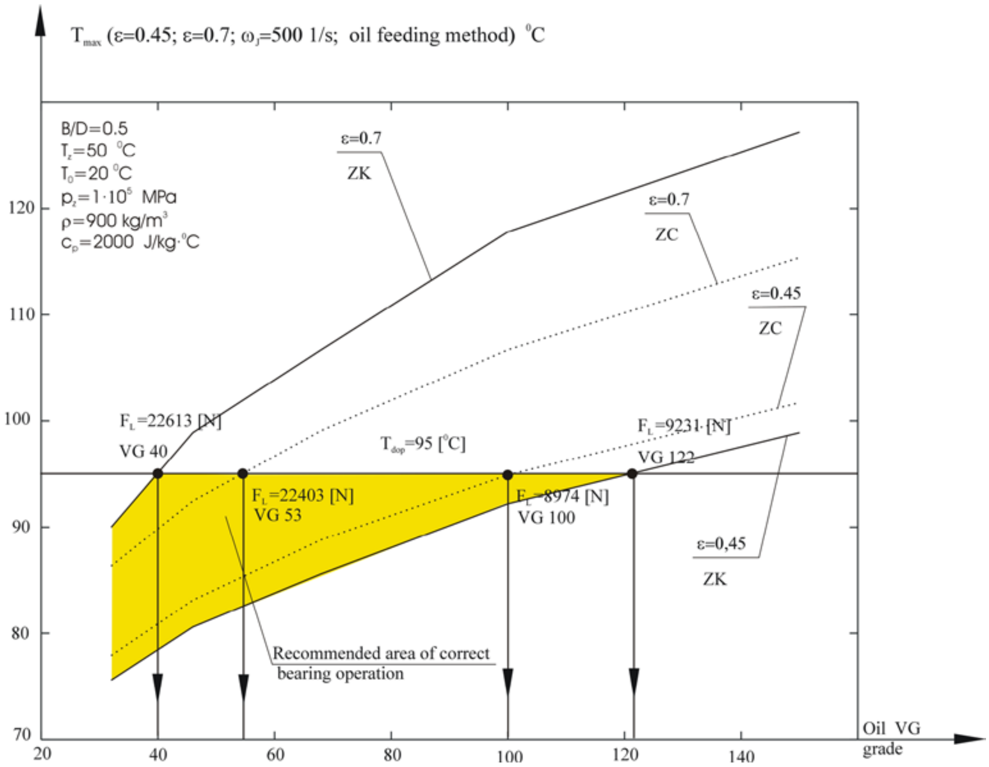


Fig. 7. The effect of oil grade on the maximum temperature in bearing. Symbols: ZC — fresh oil feeding from the bearing face side; ZK — feeding with fresh oil from a lubrication pocket

Based on research results presented in Figure 7 it can be claimed that for the eccentricity value $\varepsilon = 0.7$, bearings fed from a lubrication pocket can be operated with oils of the grade $VG_{max} = 40$, whereas oils with $VG_{max} = 53$ can be used when fed from the bearing face side. For $\varepsilon = 0.45$, the limiting grade values are $VG_{max} = 122$ and $VG_{max} = 100$, respectively.

The presented results represent the outcome of the first stage of a wider research project. In the next step, the effect of tolerance of oil operating properties on the bearing node operating parameters and dynamical properties of the bearings will be examined. The parameters influencing the lubricating properties will be also researched in the future.

References

1. Abdollahzadeh Jamalabadi MY, Alamian R, Yan W-M, Li LKB, Leveneur S, Safdari Shadloo M. Effects of Nanoparticle Enhanced Lubricant Films in Thermal Design of Plain Journal Bearings at High Reynolds Numbers. Symmetry 2019; 11(11): 1353, https://doi.org/10.3390/sym11111353.
2. Allmaier H, Priestner C, Reich F M, Priebsch H H, Novotny-Farkas F. Predicting friction reliably and accurately in journal bearings-extending the EHD simulation model to TEHD. Tribology International 2013; 58: 20-28, https://doi.org/10.1016/j.triboint.2012.08.015.
3. Blaut J, Breńkacz Ł. Application of the Teager-Kaiser energy operator in diagnostics of a hydrodynamic bearing. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (4): 757-765, https://doi.org/10.17531/ein.2020.4.20.
4. Budzik G, Mazurkow A. Modelling and Testing of Dynamic Properties of C0-45 Turbochargers. Scientific Journal of Silesian University of Technology , Series Transport 2017; 97: 17-25, https://doi.org/10.20858/sjsutst.2017.97.2
5. Dang PV, Chatterton S, Pennacchi P. Static Characteristics of a Tilting Five-Pad Journal Bearing with an Asymmetric Geometry. Actuators 2020; 9(3): 89, https://doi.org/10.3390/act9030089.
6. Dyk S, Smolik L, Rendl J. Predictive capability of various linearization approaches for floating-ring bearings in nonlinear dynamics of turbochargers. Mechanism and Machine Theory 2020; 149: 103843, https://doi.org/10.1016/j.mechmachtheory.2020.103843.
7. Feng H, Jiang S, Ji A. Investigations of the static and dynamic characteristics of water-lubricated hydrodynamic journal bearing considering turbulent, thermohydrodynamic and misaligned effects. Tribology International 2019; 130: 245-260, https://doi.org/10.1016/j.triboint.2018.09.007.
8. Garcia M, Bou-Said B, Rocchi J, Grau G. Refrigerant foil bearing behavior-A Thermo-HydroDynamic study (Application to rigid bearings). Tribology International 2013; 65: 363-369, https://doi.org/10.1016/j.triboint.2012.12.006.
9. Kuznetsov E, Glavatskih S, Fillon M. THD analysis of compliant journal bearings considering liner deformation. Tribology International 2011; 44: 1629-1641, https://doi.org/10.1016/j.triboint.2011.05.013.
10. Li B, Sun J, Zhu S, Fu Y, Zhao X, Wang H, Teng Q, Ren Y, Li Y, Zhu G. Thermohydrodynamic lubrication analysis of misaligned journal bearing considering the axial movement of journal. Tribology International 2019; 135: 397-407, https://doi.org/10.1016/j.triboint.2019.03.031.
11. Li Y, liang F, Zhou Y, Ding S, Du F, Zhou M, Bi J, Cai Y. Numerical and experimental investigation on thermohydrodynamic performance of turbocharger rotor-bearing system. Applied Thermal Engineering 2017; 121: 27-38, https://doi.org/10.1016/j.applthermaleng.2017.04.041.
12. Mazurkow A. The design method of slide bearings with the floating ring bearing. Lap Lambert Academic Publishing, Saarbruecken, Germany, 2013, ISBN: 978-3-659-37505-7.
13. Mazurkow A. The study of journal slide bearings start-up with hydrodynamic lubrication. 55 Tribologie- Fachtagung, Gesellschaft fuer Tribologie e.V. Goettingen, 2014, ISBN: 978-3-00-046545-1.
14. Nikolic N, Antonic Z, Doric J, Ruzic D, Galambos S, Jocanovic M, Karanovic V. An Analytical Method for the Determination of Temperature Distribution in Short Journal Bearing Oil Film. Symmetry 2020; 12(4): 539, https://doi.org/10.3390/sym12040539.
15. Ramos D J, Daniel G B. A new concept of active hydrodynamic bearing for application in rotating systems. Tribology International 2021; 153: 106592, https://doi.org/10.1016/j.triboint.2020.106592.
16. Sadabadi H, Sanati Nezhad A. Nanofluids for Performance Improvement of Heavy Machinery Journal Bearings: A Simulation Study. Nanomaterials 2020; 10(11): 2120, https://doi.org/10.3390/nano10112120.
17. Santos N D S A, Roso V R, Faria M T C. Review of engine journal bearing tribology in start-stop applications. Engineering Failure Analysis 2020; 108: 104344, https://doi.org/10.1016/j.engfailanal.2019.104344.
18. Strzelecki S, Kusmierz L, Ponieważ G. Thermal deformation of pads in tilting 5-pad journal bearing. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2008; 2: 12-16.
19. Suh J, Choi YS. Pivot design and angular misalignment effects on tilting pad journal bearing characteristics: Four pads for load on pad configuration. Tribology International 2016; 102: 580-599, https://doi.org/10.1016/j.triboint.2016.05.049.
20. Tian L, Wang W J, Peng Z J. Dynamic behaviours of a full floating ring bearing supportedturbocharger rotor with engine excitation. Journal of Sound and Vibration 2011; 330: 4851-4874, https://doi.org/10.1016/j.jsv.2011.04.031.
21. Wang Y, Fang X, Zhang C, Cchen X, Lu J. Lifetime prediction of self-lubricating spherical plain bearings based on physics-of-failure model and accelerated degradation test. Eksploatacja i niezawodnosc - Maintenance and Reliability 2016; 18 (4): 528-538, http://dx.doi.org/10.17531/ein.2016.4.7.
22. Wodtke M, Litwin W. Water-lubricated stern tube bearing - experimental and theoretical investigations of thermal effects. Tribology International 2021; 153: 106608, https://doi.org/10.1016/j.triboint.2020.106608.
23. Xie Z, Shen N, Zhu W, Tian W, Hao L. Theoretical and experimental investigation on the influences of misalignment on the lubrication performances and lubrication regimes transition of water lubricated bearing. Mechanical Systems and Signal Processing 2021; 149: 107211, https://doi.org/10.1016/j.ymssp.2020.107211.
24. Xie Z, Zhang Y, Zhou J, Zhu W. Theoretical and experimental research on the micro interface lubrication regime of water lubricated bearing. Mechanical Systems and Signal Processing 2021; 151: 107422, https://doi.org/10.1016/j.ymssp.2020.107422.
25. Zhang Y, Yang L, li Z, Yu L. Research on static performance of hydrodynamically lubricated thrust slider bearing based on periodic harmonic. Tribology International 2016; 95: 236-244, https://doi.org/10.1016/j.triboint.2015.11.026.
26. Zhou HL, Feng GQ, Luo GH, Ai YT, Sun D. The dynamic characteristics of a rotor supported on ball bearings with different floating ring squeeze film dampers. Mechanism and Machine Theory 2014; 80: 200-213, https://doi.org/10.1016/j.mechmachtheory.2014.04.016.
27. DIN 31652, Part 1, 2, 3. Plain bearings - Hydrodynamic plain journal bearings under steady-state conditions.
28. ISO 1940-1:2003 Mechanical vibration - Balance quality requirements for rotors in a constant (rigid) state - Part 1: Specification and verification of balance tolerances.
29. ISO 3448:1992 Industrial liquid lubricants - ISO viscosity classification.

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 23 (2021), Issue 2
journal homepage: http://www.ein.org.pl

Article citation info:
Chen C, Wang C, Lu N, Jiang B, Xing Y. A data-driven predictive maintenance strategy based on accurate failure prognostics. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 387–394, http://doi.org/10.17531/ein.2021.2.19.

A data-driven predictive maintenance strategy based on accurate failure prognostics

Indexed by: Web of Science Group

Chuang Chen[a], Cunsong Wang[a], Ningyun Lu[a,*], Bin Jiang[a], Yin Xing[b]

[a] College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
[b] School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China

Highlights

- Degradation feature selection module helps to lessen calculation burden.
- The prognostic model provides degradation trends online for failure prognostics.
- The perfect time for taking maintenance activities can be determined.

Abstract

Maintenance is fundamental to ensure the safety, reliability and availability of engineering systems, and predictive maintenance is the leading one in maintenance technology. This paper aims to develop a novel data-driven predictive maintenance strategy that can make appropriate maintenance decisions for repairable complex engineering systems. The proposed strategy includes degradation feature selection and degradation prognostic modeling modules to achieve accurate failure prognostics. For maintenance decision-making, the perfect time for taking maintenance activities is determined by evaluating the maintenance cost online that has taken into account of the failure prognostic results of performance degradation. The feasibility and effectiveness of the proposed strategy is confirmed using the NASA data set of aero-engines. Results show that the proposed strategy outperforms the two benchmark maintenance strategies: classical periodic maintenance and emerging dynamic predictive maintenance.

Keywords

predictive maintenance, failure prognostics, performance degradation, maintenance cost.

1. Introduction

Many of modern engineering systems operate in highly demanding environments. During long-term continuous operation under extreme conditions, operation performance inevitably deteriorates over time [1]. When reaching a critical degradation degree, underperformed components or subsystems might fail and risk the system safety [7]. Well-timed maintenance is a core desire in all engineering systems.

Maintenance strategies can be categorized into two types: preventive maintenance and corrective maintenance [8]. Preventive maintenance schedules proactive maintenance activities routinely; while corrective maintenance is an unscheduled strategy that attempts to restore the system after failures [4]. For those systems that have excessive demands on safety and reliability, preventive maintenance is the main stream. Traditional preventive maintenance is based on the serving time and the probability distribution of trouble-free operation time span of the system. So, it is also termed as time-based maintenance (TBM). Its conservation is obvious. On one hand, taking intensive preventive maintenance results in excessive maintenance; and on the other hand, preventive maintenance with fixed time span can't avoid unexpected faults or the faults with insufficient prior knowledge [10]. To improve cost-effectiveness ratio of preventive maintenance, condition based maintenance (CBM) that takes into account the actual operating conditions of the system over time, has been proposed and received considerable attentions from academia to industry over the last decade [19].

In the existing CBM strategies, degrading system condition is often described by stochastic modeling, such as a Markov chain with multiple discrete states [13, 14, 15, 16] or a stochastic process model with a continuous degradation state [5, 6, 22]. These stochastic-model-based CBM strategies either require that the transition probabilities of system states are known in advance or can be learned from the historical reliability data, or require that there exists a stochastic process characterizing the system degradation mechanism. However, in practice, it is difficult or even impossible to obtain the accurate probability distributions of all possible transitions of system states and the accurate degradation mechanism of a complex engineering system with affordable cost. To avoid these tough problems of the existing stochastic-model-based CBM strategies, in recent years, machine learning based methods that can be independent of the system degradation mechanism are applied to the field of prognostics and health management (PHM) [18]. In this emerging field, a trend of maintenance technology is to make maintenance decision based on multivariate condition monitoring and failure prognostics [2]. For example, a new deep neural network structure called long short-term memory (LSTM) network was used to discover the underlying time series patterns for predicting the

(*) Corresponding author.
E-mail addresses: C. Chen - chenchuang@nuaa.edu.cn, C. Wang - wangcunsong@nuaa.edu.cn, N. Lu - luningyun@nuaa.edu.cn, B. Jiang - binjiang@nuaa.edu.cn, Y. Xing - xingyin@hhu.edu.cn

system remaining useful life (RUL) [21]. In paper [3], the authors adopted a restricted Boltzmann machine to pre-train the abstract features for LSTM input. Moreover, a two-dimensional grid LSTM is designed to improve the prediction accuracy of fuel cell performance degradation [9].

The above machine learning based research only focuses on life prediction, and does not consider the maintenance decision-making issues. Recently, a novel dynamic predictive maintenance (PdM) framework using LSTM network for failure prognostics has been developed [12]. The authors have discussed in detail the advantages of PdM over other maintenance strategies, and a complete framework from data-driven prognostics to maintenance decisions has been given. In our past work, an effective data-driven degradation prognostic technique has been developed with good verification results for the aero-engine system [20]. The work of this paper is a follow-up of [12] and [20], and the main contribution is to develop a data-driven PdM maintenance strategy to make long-term, reliable maintenance decisions for engineering systems. In detail, we design a module of degradation feature selection. It can enable the failure prognostics and maintenance decision-making to have lower computing load, faster convergence speed and better robustness in presence of uncertainties. More accurate failure prognostics can be realized via the multivariate LSTM network whose inputs are the selected degradation features. The prognostic model can provide the future degradation trend online for failure prognosis. For maintenance decision-making, the perfect time for taking maintenance activities can be determined by evaluating the maintenance cost online based on the failure prognostic results of performance degradation. Correspondingly, long-term, reliable maintenance decisions can be realized, which is crucial for planning maintenance, inventory and production activities in advance.

The remainder of this paper is organized as follows. In Section 2, an enhanced data-driven PdM strategy is presented, including implementation details and performance evaluation, under the framework of [12]. In Section 3, the feasibility and effectiveness of the proposed PdM strategy will be confirmed using the NASA data set of aero-engines. Conclusions and future works will be discussed in Section 4.

2. An enhanced data-driven PdM strategy

2.1. Key idea

A novel data-driven dynamic PdM framework has been proposed in [12], which has provided a complete process from data-driven prognostics to maintenance decisions. The entire process, as shown in Fig. 1, functionally includes three parts: LSTM modeling, online failure prognosis and maintenance decisions.

The LSTM step includes training of an LSTM classifier and using the LSTM classifier to determine the degradation label of online measurements. It deals with the multivariate raw data directly and all data are used as the inputs of LSTM model. This may cause extensive computing load, low convergence speed, low robustness of the LSTM modeling, and ultimately reduce the accuracy of failure prognosis. Also, the LSTM network only provides the probabilities of system failure at the current moment. This limits the decision-making to be instantaneous. Instantaneous decision-making of system only answers whether or not the system need maintenance activities at the current moment. It cannot give the exact time when the system must take preventive maintenance activities. Apparently, in practice, a long-term, reliable decision-making is more valuable for industrial organizers to plan maintenance, inventory and production activities in advance.

To overcome the above issues, this paper proposes an enhanced dynamic PdM strategy that can enable to achieve future failure prognosis and long-term, reliable maintenance decision-making. The main steps are shown in Fig. 2. Compared with the original PdM framework in Fig. 1,

(1) in data preprocessing step, the multivariate raw data are firstly preprocessed to extract the features that can reflect the degradation trends;
(2) in LSTM modeling step, an extra LSTM regression model is introduced for predicting the future degradation trends of system;
(3) in the decision-making step, the predicted failure probabilities at different moments in future are used to make long-term

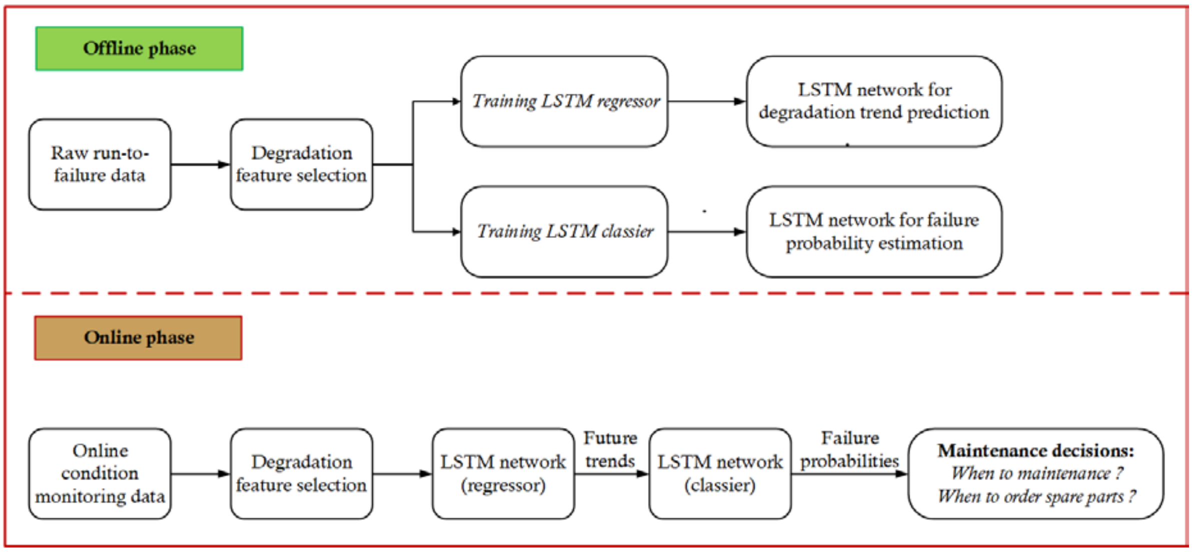


Fig. 2. Enhanced dynamic PdM framework

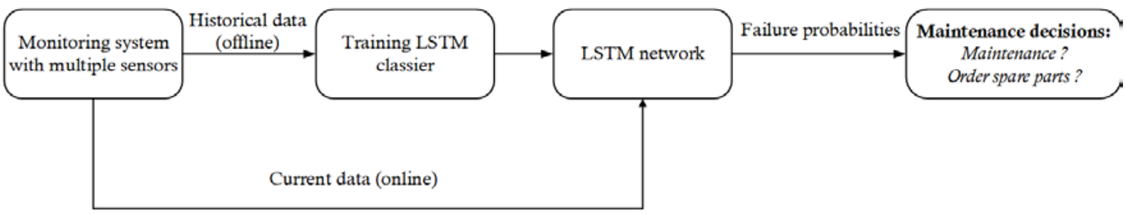


Fig. 1. The dynamic PdM framework [12]

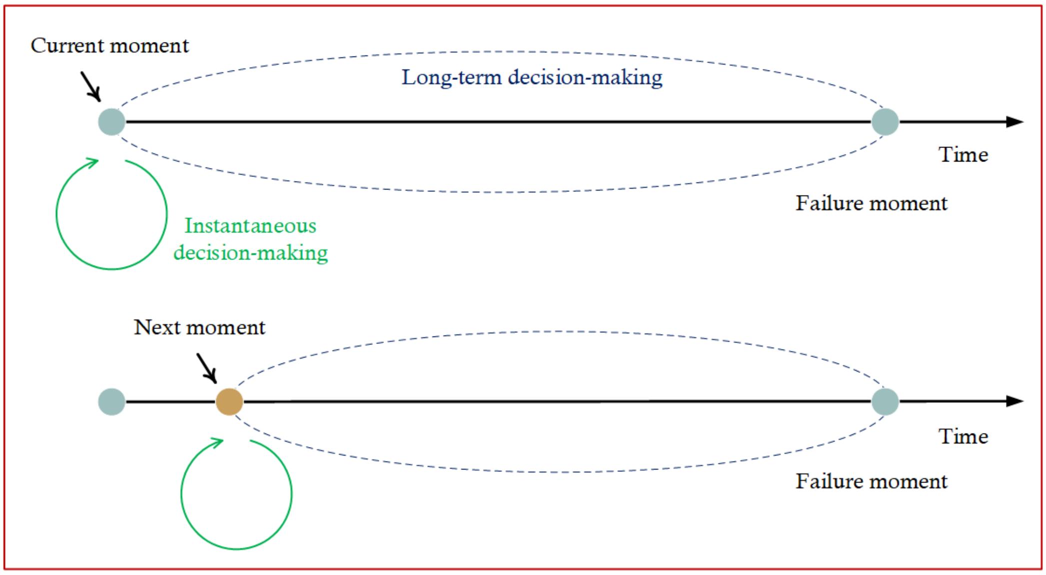


Fig. 3. Dynamic instantaneous and long-term decision-making processes

maintenance decisions, *e.g.*, to decide when the system needs taking maintenance activities and ordering the spare parts.

Fig. 3 illustrates the difference between the dynamic instantaneous and long-term decision-making processes. At the current moment, the instantaneous decision-making answers whether or not the system need maintenance activities, while the long-term decision-making gives the exact time when the system must take preventive maintenance activities. Obviously, the long-term decision-making has a broader vision. As the operation time of the system increases, the sensors will obtain more condition monitoring data, making the decision-making results more accurate.

2.2. Degradation feature selection and improved failure prognosis via LSTM

In practice, the sensor measurements are often contaminated with noises. Noises may conceal the tenuous degradation trend. So data de-noising should be conducted in the data pre-processing phase. To do so, the simple but effective moving average method is employed to extract the system degradation trends [20]. This process is briefly described as follows. Firstly, all available historical condition monitoring data can be arranged into a three-dimensional data $X(I \times J \times K)$, where I denotes the number of samples, J denotes the number of measuring variables and K denotes the operation cycle. The k th value of the j th variate in the i th sample is denoted as $x_{ij}(k)$. Thus, the degradation values using moving average are given by:

$$\tilde{x}_{ij}(k) = \sum_{h=k-n+1}^{k} x_{ij}(h) / n \text{ with } k = n, n+1, \cdots, K_i \tag{1}$$

where n is the size of moving window. Then, the Z-score normalization is used to handle the different ranges of sensor measurements. Normalized sensor measurements are given by:

$$\hat{x}_{ij}(k) = (\tilde{x}_{ij}(k) - \mu) / \delta \tag{2}$$

whereμ μ and δ denotes the mean and standard deviation of these degradation values, respectively, and are given by:

$$\mu = \sum_{k=1}^{K_i} \tilde{x}_{ij}(k) / K_i \tag{3}$$

$$\delta = \sqrt{\sum_{k=1}^{K_i} (\tilde{x}_{ij}(k) - \mu)^2 / (K_i - 1)} \tag{4}$$

In addition, eliminating usefulness data is necessary before LSTM network modeling since it can generally improve the performances of modeling, failure prognosis and decision making. Therefore, a module of degradation feature selection is included in the proposed maintenance strategy. In this paper, the correlation and trend indicators are adopted for degradation feature selection due to their effectiveness. The correlation and trend indicators are given by:

$$\rho_{ij} = 1 - 6\sum_{k=1}^{K_i} d_{\hat{x}_{ij}(k)}^2 / (K_i^3 - K_i) \tag{5}$$

$$T_j = \sum_{i=1}^{I} \left(1 \cdot \delta(\rho_{ij} > 0) + 0.5 \cdot \delta(\rho_{ij} = 0)\right) / I \tag{6}$$

where $d_{\hat{x}_{ij}(k)}$ denotes the difference between ranks for each $\hat{x}_{ij}(k)$ and k, and $\delta(x)$ is the direct function, i.e., $\delta(x) = 1$ when x is true and $\delta(x) = 0$ otherwise. According to the two indicators, the crucial features can be selected by the criterion, $|\rho_{ij}| \geq 0.5 \;\&\&\; T_j == 0 \text{ or } 1$ [20].

Algorithm 1 Degradation prognostic model based on LSTM network

Input: $\hat{X}(I \times F \times K)$

Output: A well-trained multivariate LSTM network

Process:

1: **for** $i = 1, 2, \cdots, I$ **do**

2: **for** $j = 1, 2, \cdots, F$ **do**

3: $net.input = \hat{x}_{ij}(1 : K_i - 1)$;

4: $net.output = \hat{x}_{ij}(2 : K_i)$;

5: end for

6: end for

7: # LSTM network training

8: *LSTM ← train (net.input, net.output, solver.adam, regularization.dropout)*;

9: **return** well-trained network parameters.

Next, to obtain the failure probabilities at different moments in future, a multivariate LSTM regressor for degradation trend prediction is first trained with historical data (see Algorithm 1). It is noted that, the multivariate LSTM network can exploit the nature of the evolving degradation trend [23], and in Algorithm 1, $\widehat{X}(I \times F \times K)$ denotes the pre-processed data with F important features. Fig. 4 shows a schematic diagram of the degradation trend prediction. For the online condition monitoring data (duration: 1-t), they will be pre-processed in the same way, and then fed into the well-trained multivariate LSTM regressor. The regressor can predict the degradation trends of system in future.

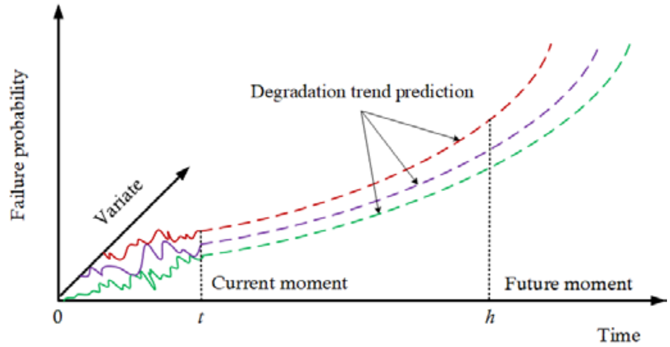


Fig. 4. Schematic diagram of degradation trend prediction

Similar to [12], a multivariate LSTM classier for failure probability estimation is trained with historical data (see Algorithm 2). It is noted that, in Algorithm 2, $R(I \times 1 \times K)$ denotes the RUL data, and the RUL value of k th cycle of the i th sample is denoted as $r_{i1}(k)$. The degradation data will be labeled by two classes: Deg1 and Deg2. Deg1 represents the case where the system RUL time is greater than or equal to the time window w_0, *i.e.*, $RUL \geq w_0$. Deg2 means $RUL < w_0$. The two labels can be regarded as two degradation states with different degrees, like allowable degradation and intolerable degradation.

Algorithm 2 Failure prognostic model based on LSTM network

Input: $\widehat{X}(I \times F \times K)$ and $R(I \times 1 \times K)$

Output: A well-trained multivariate LSTM network

Process:

1: **for** $i = 1, 2, \cdots, I$ **do**

2: **for** $j = 1, 2, \cdots, F$ **do**

3: **for** $k = 1, 2, \cdots K_i$ **do**

4: # Data labeling

5: $r_{i1}(k) \leftarrow 1 \cdot \delta(r_{i1}(k) \geq \Delta T)$;

6: $r_{i1}(k) \leftarrow 2 \cdot \delta(r_{i1}(k) < \Delta T)$;

7: end for

8: $net.input = \hat{x}_{ij}(1:K_i)$;

9: $net.output = r_{i1}(1:K_i)$;

10: end for

11: end for

12: *LSTM* ← *train* (*net.input, net.output, solver.adam, regularization.dropout*);

13: **return** well-trained network parameters.

In practice, due to technical and logistical constraints, maintenance activities cannot be carried out at anytime and anywhere. As an illustration, the maintenance activities for train or aircraft engines cannot be realized during their journeys. Maintenance activities can be performed only at the inspection moment. It is assumed that the inspection interval ΔT between two successive inspections is constant. If the RUL of the system at some inspection moment h in the future is less than ΔT, it means the system has failed at the next moment $h + \Delta T$. Hence, the time window is equal to inspection interval, *i.e.*, $w_0 = \Delta T$.

The predicted degradation trends are ultimately fed into the well-trained LSTM classier, and thus the failure probabilities at different moments in future are obtained.

2.3. Improved maintenance decision-making method

The following long-term maintenance strategy attempts to answer the exact points in the future to take maintenance activities and to order spare parts. The optimal maintenance moment can be determined by choosing the solution with the lower cost from the *expected-preventive-maintenance (PM) cost* and the *no-PM cost* based on the predicted failure probabilities.

The expected-PM cost is defined as follows. At a future moment h ($h = t + \Delta T, t + 2\Delta T, \cdots$), all the costs associated with the preventive maintenance actions such as replacing the worn parts with new ones, system cleaning and adjustment, and the inventory cost of spare parts, are summed up to be the expected-PM cost, which can be denoted as C_p. An important assumption to note here is, the system after taking the PM actions can be restored to be "as good as new" state, or in other words, perfect maintenance is considered in this paper.

If no PM actions are taken at the moment h, there will be no PM cost from the current moment t to the future moment h, but there exists the failure risk of the running system between h and $h + \Delta T$. In this case, one must consider the no-PM cost, which includes the corrective maintenance cost C_c with unexpected failures and the out-of-stock cost C_{os} in the case of unavailable spare parts. Thus, the expected cost with the decision of no-PM action is defined as $(C_c + C_{os}) \cdot P(RUL_h < \Delta T)$, where $P(RUL_h < \Delta T)$ denotes the probability of the unexpected failures between the inspection period $[h, h + \Delta T)$.

Fig. 5 shows the decision process based on the above-mentioned maintenance costs. If the expected-PM cost is lower than or equal to the no-PM cost, PM activities should be taken. Otherwise, no maintenance activity is required in the inspection period $[h, h + \Delta T)$, *i.e.*:

$$C_p \leq (C_c + C_{os}) \cdot P(RUL_h < \Delta T). \tag{7}$$

Thus, the optimal maintenance moment $t^*_{maintenance}$ can be obtained as:

$$t^*_{maintenance} = \inf_{h=t+\Delta T, t+2\Delta T, \cdots} \left\{ h : \middle| C_p \leq (C_c + C_{os}) \cdot P(RUL_h < \Delta T) \right\} \tag{8}$$

Ordering of spare parts should be implemented before the maintenance activities. If the longest advanced ordering time is L, the optimal ordering moment t^*_{order} can be given by:

$$t^*_{order} = t^*_{maintenance} - L. \tag{9}$$

2.4. Implementation and performance evaluation

With the historical condition monitoring data and the real-time condition monitoring data of the system, the optimal preventive maintenance and ordering moments are obtained online according to the following procedures:

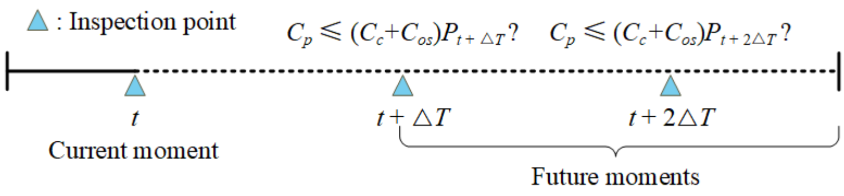


Fig. 5. Decision process based on the maintenance cost

(1) Obtain crucial degradation features according to the correlation and trend criteria;
(2) Obtain future degradation trends by feeding the crucial degradation features into the network in Algorithm 1;
(3) Obtain failure probabilities at different moments in future by feeding the predicted degradation trends into the network in Algorithm 2;
(4) Calculate the expected-PM cost and no-PM cost according to Eq. (7);
(5) Obtain optimal maintenance time $t^*_{maintenance}$ and optimal ordering moment t^*_{order} according to Eq. (8) and Eq. (9).

To evaluate the maintenance strategy, maintenance cost rate (MCR) [12] is considered. It is defined as the ratio between the total maintenance cost and the total life cycle duration. The strategy with lower MCR is considered to have better performance. It is worth noting that, there two possible scenarios in real-world maintenance activities.

If the scheduled preventive maintenance moment is ahead of the actual failure moment of the system, the preventive maintenance activities will be performed. In this case, the available spare parts can arrive in time thanks to the scheduled order moment. Correspondingly, the MCR with no system failure (denoted by MCR_p) is given by:

$$MCR_p = \frac{C_p}{t^*_{maintenance}} \tag{10}$$

Contrarily, if the system is failed before the scheduled preventive maintenance moment, the corrective maintenance has to be taken. In this case, there is no available spare parts, and the corrective maintenance cost C_c and the out-of-stock cost C_{os} with unavailable spare parts have to be paid. Thus, the MCR with system failure (denoted by MCR_c) is given by:

$$MCR_p = \frac{C_c + C_{os}}{[T_F / \Delta T]^+ \cdot \Delta T} \tag{11}$$

where T_F denotes the actual failure moment of the system, and $[x]^+$ means taking a smallest integer more than or equal to a real number x.

3. Case study

3.1. Data description

To verify the feasibility and effectiveness of the proposed maintenance strategy, the Turbofan Engine Degradation Simulation Data Set [11] provided by NASA Ames Prognostics Data Repository is referred. The data set is generated by C-MAPSS tool that simulates the degradation process of the main components of turbofan engines, *e.g.*, fan, low-pressure compressor (LPC), high-pressure compressor (HPC), high pressure turbine (HPT) and low pressure turbine (LPT). Twenty-one sensors are installed inside the engine for monitoring the conditions of the engine. The first nine sets of data are obtained by direct measurement of sensors #1~#9, while the remaining data are gained by soft measurement of sensors #10~#21 [17].

In the experiment, the available data set "FD001" that describes the gradual degradation process of HPC under a constant work condition is selected to show the use of the proposed maintenance strategy. The data set contains the "train_FD001.txt" composed of 100 complete run-to-failure data $X(100 \times 21 \times K)$ $(127 \le K \le 362)$, the "test_FD001.txt" composed of 100 incomplete run-to-failure data $X'(100 \times 21 \times K')$ $(31 \le K' \le 303)$ and the "RUL_FD001.txt" providing the actual RUL information.

3.2. Offline modeling

Fig. 6. shows the parts of results of degradation feature selection. For sensor #1 (see Fig. 6(a)), its correlation indicator in each engine training sample is always 0, which means that the monitoring variable remains constant during the engine operation phase. Obviously, such monitoring variable has no effect on the system failure prognosis and should be eliminated. For sensor #4 (see Fig. 6(b)), its correlation indicator in each training sample is always greater than 0.5. This means that such monitoring variable has been positively correlated with operating time (flight cycle). In addition, its trend indicator value is 1, indicating that it has a monotonous upward trend. Thus, the sensor #4 are retained. Regarding the sensors # 9 and 13 (see Fig. 6(c) and (see Fig. 6(d))), they are also not proper degradation features since their correlation indicators are not still positive or negative. Finally, only seven sensors are selected, *i.e.*, the sensors #4, #7, #11, #12, #15, #20 and #21. After some experiments, the value 20 is taken as the moving window size due to the best performance on test data set. Then, the data are normalized using the Z-score method (see Eq. (2)) so that they have the same means and variances.

With reduced degradation feature data, the next step is to train degradation prognostic model and failure prognostic model using LSTM neworks. Notably, the degradation prognostic model is used to obtain the evolving degradation trends, while the failure prognostic model is used to obtain the failure probabilities at different moments in future based on the predicted degradation trends. In the LSTM network, the number of iterations is set to 50, the dropout rate is set to 0.2, the number of 1st LSTM units is set to 100 and the number of 2nd LSTM units is set to 50 [12]. Using **Algorithm 1**, the degradation prognostic model is built. Fig. 7 shows the offline degradation trend prediction results for training Engines #1, #2 and #3. It can be seen that regardless of Engines #1, #2 or #3, the offline predicted degradation trend values are very close to the actual degradation trend values. The offline training root-mean-square errors (RMSEs) of three engines are 0.50, 0.43 and 0.47, respectively, which indicates that the degradation prognostic model has been well built.

Given the inspection interval $\Delta T = 10$, the failure prognostic model can be built based on **Algorithm 2**. Fig. 8 shows the offline failure probability estimation results for training Engines #1, #2 and #3. The abscissa represents the operation cycle of the engine, while the ordinate "1" and "2" represent two categories: Deg1 and Deg2, respectively.

For the training Engine #1, the predicted cycles of label "2" are 1-185 cycles whose corresponding probabilities satisfy $P(RUL < 10) < 0.5$, while the actual cycles are 1-183 cycles. With regard to the training Engine #2 and #3, the predicted cycles of label "2" are 1-277 and 1-173 cycles, while the actual cycles are 1-278 and 1-170 cycles, respectively. These results shows that the failure prognostic model has been well built.

3.3. Online maintenance scheduling

As an example, the testing Engine #1 is used to illustrate the online prognostics. The online prognostics contain the online degradation trend prediction and online failure probability estimation. Fig. 9 shows the online trend prediction results for testing Engine #1. The condition monitoring data collected up to present are 31 cycles for the

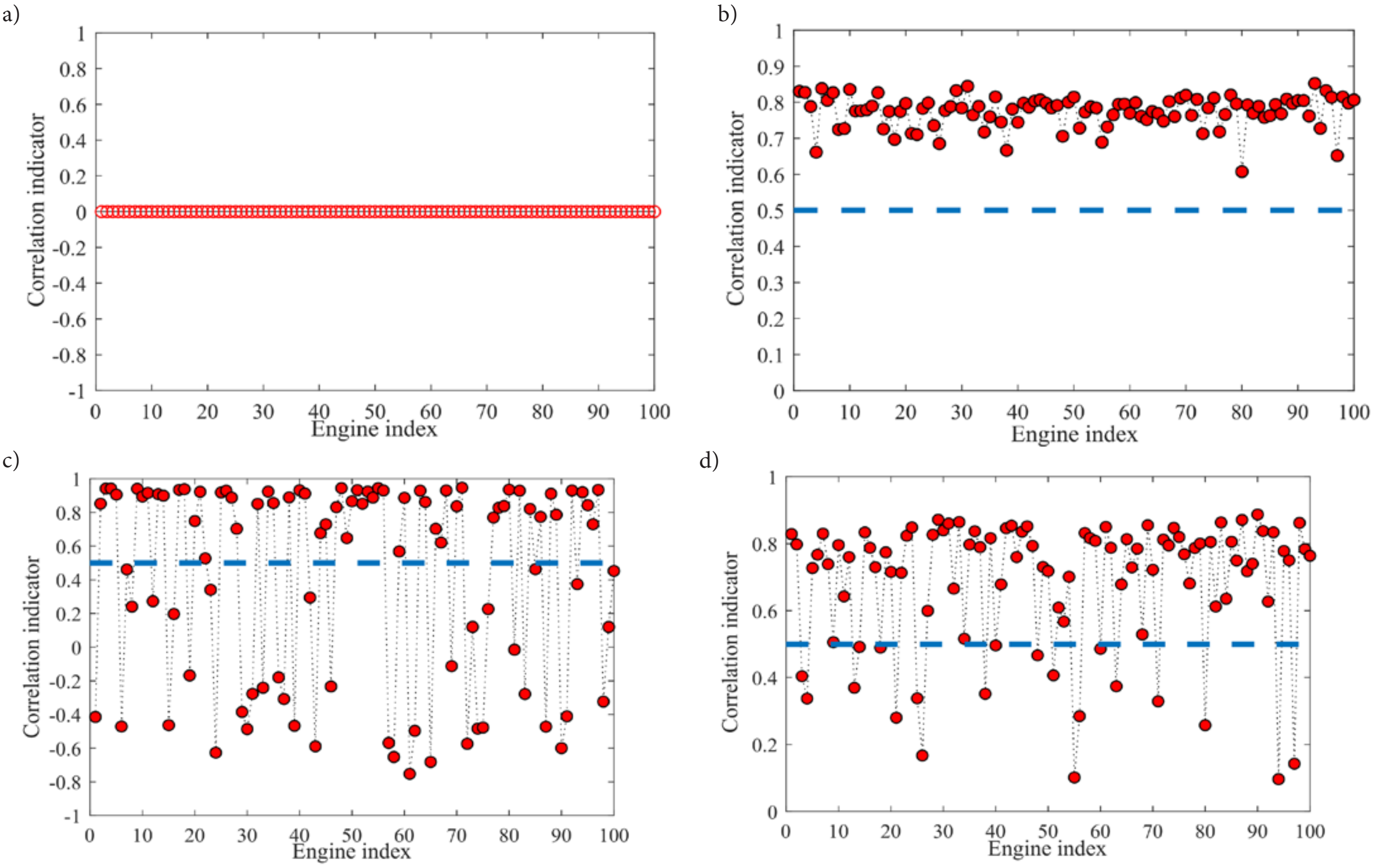


Fig. 6. Parts of results of degradation feature selection: a) Sensor #1, b) Sensor #4, c) Sensor #9, d) Sensor #13

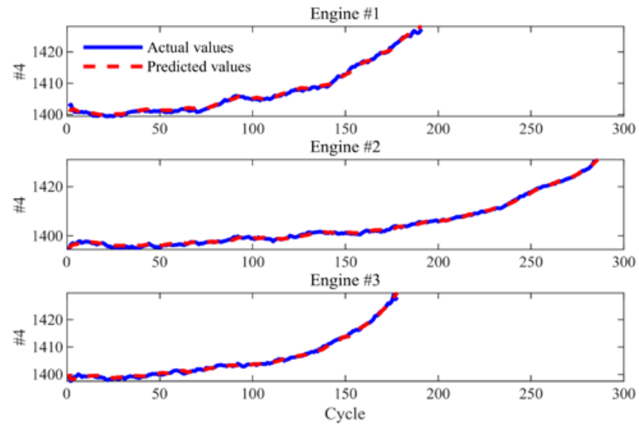


Fig. 7. Offline degradation trend prediction results for training Engines #1, #2 and #3

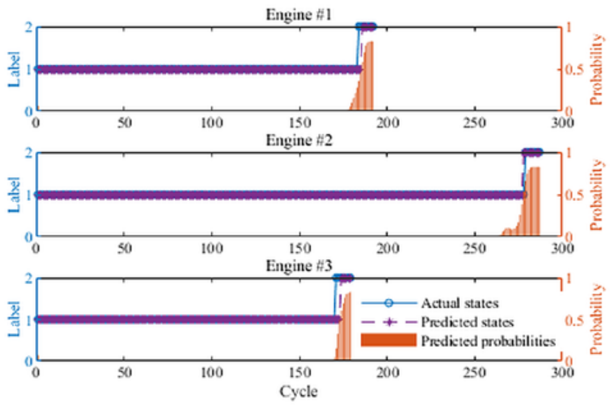


Fig. 8. Offline failure probability estimation results for training Engines #1, #2 and #3

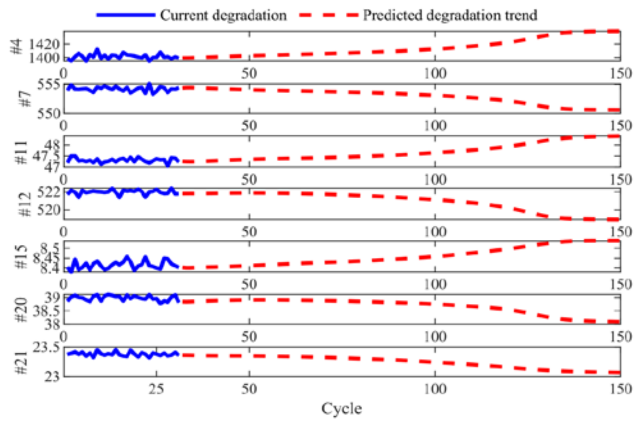


Fig. 9. Online trend prediction results for testing Engine #1

testing Engine #1. It can be seen that the conditions of the engine are gradually deteriorating over time.

Next, these predicted trend values are fed into the well-trained failure prognostic model. Fig. 10 shows the online failure probability estimation results for testing Engine #1. It can be seen that as the operation cycle of the engine increases, the failure probability increases. When the operation cycle exceeds the Cycle 133, the failure probabilities are stable with a high value (0.8278). Note that the moment that the first predicted failure probability crosses 0.5 is Cycle 128, indicating that the RUL of the engine will only survive for 10 days. Thus, the estimated end of life (EOL) of testing Engine #1 is Cycle 138, while the actual EOL is Cycle 143 according to the "RUL_FD001.txt". This indicates the failure prognostic is accurate.

Suppose that the preventive maintenance cost $C_p = 100$, the corrective maintenance cost $C_c = 500$ and the out-of-stock cost $C_{os} = 10$

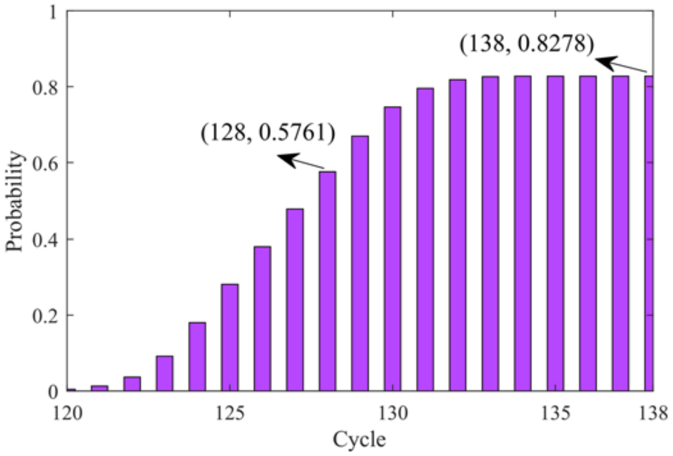


Fig. 10. Online failure probability estimation results for testing Engine #1

of the aero-engine. According to Eq. (7), the expected-PM cost and no-PM cost can be calculated, as shown in Table 1.

Table 1. Results of the Expected-PM costs and no-PM costs

Operating cycle	Failure probability	PM-cost	No-PM cost
31	0	100	0
32	0	100	0
⋮	⋮	⋮	⋮
122	0	100	0
123	0.0018	100	0.9180
124	0.0046	100	2.3460
125	0.0128	100	6.5280
126	0.0363	100	18.5130
127	0.0916	100	46.7160
128	0.1801	100	91.8510
129	0.2812	100	143.4120

Before the 129th cycle, the expected-PM cost is higher than the no-PM cost, while in the 129th cycle, the expected-PM cost is lower than the expected no-PM cost. Hence, theoretically, the optimal maintenance moment is the 129th cycle. However, in practice, the maintenance activities can be carried out only at the inspection moments, so the real maintenance activities will be taken at the 120th cycle. If the logistic service department can provide the lead time of 20 cycles in ordering the spare parts, the optimal order moment will be 100th cycle.

3.4. Comparative results and discussion

In this section, the proposed maintenance strategy is compared with the three benchmark maintenance strategies [12]: original dynamic PdM strategy, classical periodic maintenance (PeM) strategy and ideal predicted maintenance (IPM) strategy. It is noted that, the original dynamic PdM strategy focuses on the instantaneous decision-making, while the PeM and IPM strategies can handle the long-term decision-making problem.

Firstly, the original dynamic PdM strategy is compared with the enhanced one. Table 2 lists the decision-making results of the original PdM and enhanced PdM. As for the PdM strategy presented in [12], the decision-making results are that no maintenance and no ordering of spare parts are carried out in Cycle 31 (current cycle). Obviously, this strategy provides an instant decision. Regarding the enhanced PdM strategy (the method of this paper), the scheduled maintenance time is Cycle 100 and the ordering time of spare parts is Cycle 120.

Table 2. Decision-making results via the original PdM and the enhanced PdM

Maintenance strategy	Maintenance decisions	
	Order	Maintenance
Original PdM strategy	Do not order spare parts in Cycle 31 (current cycle)	Do not maintenance in Cycle 31 (current cycle)
Enhanced PdM strategy	Go to order spare parts in Cycle 100 ($T_F = 143$)	Go to maintenance in Cycle 120 ($T_F = 143$)

As far as the failure time of Cycle 143 is concerned, the planned maintenance time and ordering time of spare parts is reasonable. It is self-evident that, the enhanced PdM strategy gives the exact time when the system must take preventive maintenance activities, which helps to plan inventory and production activities in advance.

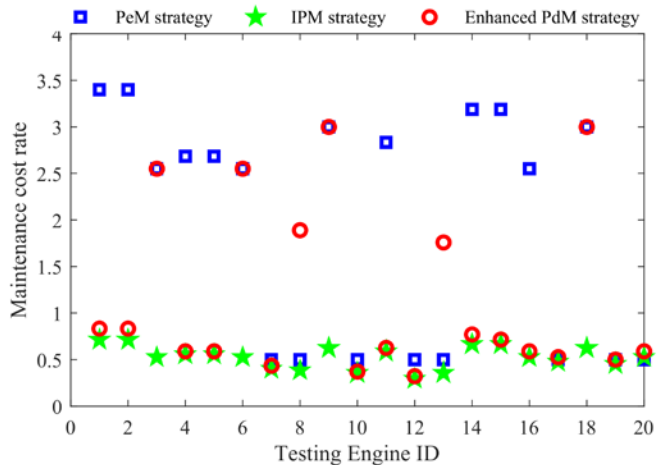


Fig. 11. Maintenance cost rates of three maintenance strategies for testing Engines #1-20

Secondly, the PeM strategy and the IPM strategy are compared with the proposed strategy. Considering that the PeM and IPM strategies are also aimed at the long-term decision-making, we uses the maintenance cost rate (MCR) presented in Section 2.4 to illustrate the superiority of the proposed strategy. The testing Engines #1-20 are taken as an example. Fig. 11 shows the MCRs of three maintenance strategies for testing Engines #1-20. From the 20 engine instances, the performance of the proposed maintenance strategy is highlighted. Specifically, compared with the PeM strategy, the MCRs of the proposed maintenance strategy are lower in most engine instances. This can be explained by the fact that, to ensure the engine safety, the PeM strategy is relatively conservative, resulting in excessive maintenances and poor economic efficiency. As for the IPM strategy, perfect prediction information is only an ideal hypothesis that cannot be attained in practice. From the figure, the MCRs of the proposed maintenance strategy are close to that of IPM strategy with perfect predictions. More specifically, the average MCRs of the three maintenance strategies are respectively calculated as follows: 1.9513 for the PeM strategy, 1.1515 for the enhanced PdM strategy, and 0.5270 for the IPM strategy. These results show that the proposed enhanced PdM strategy works well, allowing significantly reducing the maintenance cost rate.

4. Conclusions

As an important input of maintenance activities, the precision of failure prognosis directly affects the effectiveness of maintenance strategy formulation. Therefore, from the perspective of engineering applications, the data based failure prognosis needs to be considered jointly with maintenance decision-making to ensure the system safety and reliability. In this work, an enhanced data-driven predictive maintenance strategy has been developed. It provides a complete solution from failure prognosis to maintenance decision-making. The proposed strategy can obtain effective features reflecting the degradation trends. Also, it can achieve accurate failure prognostics and provide the failure probabilities at different moments in future. In particular, the proposed strategy solves the instantaneous decision-making problem and gives the exact time when the system must take preventive maintenance activities.

The verification results using NASA data repository reveal the feasibility and effectiveness of the proposed maintenance strategy. The performance of the proposed strategy is highlighted when compared with the decision-making results of the emerging dynamic predictive maintenance, the classical periodic maintenance and the ideal predicted maintenance. However, one limitation of the proposed strategy is, only the perfect maintenance is considered. Further work will focus on the investigation of imperfect maintenance with different levels. Also, the ambition is to develop flexible maintenance strategies by estimating the residence time of different health states.

Acknowledgement

This work was support by Natural Science Foundation of China (61873122, 62020106003) and China Scholarship Council (202006830060).

References

1. Chen C, Lu N, Jiang B, Wang C. A risk-averse remaining useful life estimation for predictive maintenance. IEEE/CAA Journal of Automatica Sinica 2021; 8(2): 412-422, https://doi.org/10.1109/JAS.2021.1003835.
2. Do P, Voisin A, Levrat E, Iung B. A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. Reliability Engineering & System Safety 2015; 133: 22-32, https://doi.org/10.1016/j.ress.2014.08.011.
3. Ellefsen AL, Bjorlykhaug E, Aesoy V, Ushakov S, Zhang H. Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture. Reliability Engineering & System Safety 2019; 183: 240-251, https://doi.org/10.1016/j.ress.2018.11.027.
4. Emami-Mehrgani B, Neumann WP, Nadeau S, Bazrafshan M. Considering human error in optimizing production and corrective and preventive maintenance policies for manufacturing systems. Applied Mathematical Modelling 2016; 40(3): 2056-2074, https://doi.org/10.1016/j.apm.2015.08.013.
5. Fan H, Hu C, Chen M, Zhou D. Cooperative predictive maintenance of repairable systems with dependent failure modes and resource constraint. IEEE Transactions on Reliability 2011; 60(1): 144-157, https://doi.org/10.1109/TR.2011.2104432.
6. Huynh KT, Barros A, Bérenguer C. Multi-level decision-making for the predictive maintenance of k-out-of-n: F deteriorating systems. IEEE Transactions on Reliability 2015; 64(1): 94-117, https://doi.org/10.1109/TR.2014.2337791.
7. Jasiulewicz-Kaczmarek M, Antosz K. Żywica P, Mazurkiewicz D, Sun B, Ren Y. Framework of machine criticality assessment with criteria interactions. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23 (2): 207-220, https://doi.org/10.17531/ein.2021.2.1.
8. Kozłowski E, Mazurkiewicz D, Żabiński T, Prucnal S, Sęp J. Machining sensor data management for operation-level predictive model. Expert Systems with Applications 2020; 159: 113600, https://doi.org/10.1016/j.eswa.2020.113600.
9. Ma R, Yang T, Breaz E, Li Z, Briois P, Gao F. Data-driven proton exchange membrane fuel cell degradation predication through deep learning method. Applied Energy 2018; 231:102-115, https://doi.org/10.1016/j.apenergy.2018.09.111
10. Mahdavi M, Monsef H, Romero R. Reliability effects of maintenance on TNEP considering preventive and corrective repairs. IEEE Transactions on Power Systems 2016; 32(5): 3768-3781, https://doi.org/10.1109/TPWRS.2016.2640178.
11. NASA. (Jun. 2020). Prognostic Data Repository. [Online]. Available: https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/
12. Nguyen KTP, Medjaher K. A new dynamic predictive maintenance framework using deep learning for failure prognostics. Reliability Engineering & System Safety 2019; 188: 251-262, https://doi.org/10.1016/j.ress.2019.03.018.
13. Nguyen KTP, Yeung T, Castanier B. Acquisition of new technology information for maintenance and replacement policies. International Journal of Production Research 2017; 55(8): 2212-2231, https://doi.org/10.1080/00207543.2016.1229069.
14. Nguyen TPK, Castanier B, Yeung TG. Maintaining a system subject to uncertain technological evolution. Reliability Engineering & System Safety 2014; 128: 56-65, https://doi.org/10.1016/j.ress.2014.04.004.
15. Papakonstantinou KG, Shinozuka M. Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part I: Theory. Reliability Engineering & System Safety 2014; 130: 202-213, https://doi.org/10.1016/j.ress.2014.04.005.
16. Papakonstantinou KG, Shinozuka M. Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation. Reliability Engineering & System Safety 2014; 130: 214-224, https://doi.org/10.1016/j.ress.2014.04.006.
17. Saxena A, Goebel K, Simon D, Eklund N. Damage propagation modeling for aircraft engine run-to-failure simulation. In Proceedings of 2008 International Conference on Prognostics & Health Management, Denver, CO, USA, 2008: 1-9, https://doi.org/10.1109/PHM.2008.4711414.
18. Sheppard JW, Kaufman MA, Wilmering TJ. IEEE standards for prognostics and health management. In Proceedings of 2008 IEEE Autotestcon, Salt Lake City, UT, USA, 2008: 97-103, https://doi.org/10.1109/AUTEST.2008.4662592.
19. Tahan M, Tsoutsanis E, Muhammad M, Karim ZA. Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review. Applied Energy 2017; 198: 122-144, https://doi.org/10.1016/j.apenergy.2017.04.048.
20. Wang C, Lu N, Cheng Y, Jiang B. A data-driven aero-engine degradation prognostic strategy. IEEE Transactions on Cybernetics 2021; 51(3): 1531-1541, https://doi.org/10.1109/TCYB.2019.2938244.
21. Wu Y, Yuan M, Dong S, Lin L, Liu Y. Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. Neurocomputing 2018; 275:167-179, https://doi.org/10.1016/j.neucom.2017.05.063.
22. You MY, Liu F, Wang W, Meng G. Statistically planned and individually improved predictive maintenance management for continuously monitored degrading systems. IEEE Transactions on Reliability 2010; 59(4): 744-753, https://doi.org/10.1109/TR.2010.2085572.
23. Zhang J, Wang P, Yan R, Gao RX. Long short-term memory for machine remaining life prediction. Journal of Manufacturing Systems 2018; 48: 78-86, https://doi.org/10.1016/j.jmsy.2018.05.011.

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 23 (2021), Issue 2

journal homepage: http://www.ein.org.pl

Article citation info:
Kumar S, Kumar P, Kumar G. Degradation assessment of bearing based on machine learning classification matrix. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2021; 23 (2): 395–404, http://doi.org/10.17531/ein.2021.2.20.

Degradation assessment of bearing based on machine learning classification matrix

Indexed by:
Web of Science Group

Satish Kumar[a], Paras Kumar[a], Girish Kumar[a,*]

[a] Delhi Technological University, Delhi, India

Highlights

- Machine learning classification matrix is used to model the degraded behavior of bearing
- Prior state of art considers the various diagnostic and prognostic model of bearing
- Classification model is developed to assess the degradation of bearing.
- The analysis results show that the percentage of accuracy of different models

Abstract

In the broad framework of degradation assessment of bearing, the final objectives of bearing condition monitoring is to evaluate different degradation states and to estimate the quantitative analysis of degree of performance degradation. Machine learning classification matrices have been used to train models based on health data and real time feedback. Diagnostic and prognostic models based on data driven perspective have been used in the prior research work to improve the bearing degradation assessment. Industry 4.0 has required the research in advanced diagnostic and prognostic algorithm to enhance the accuracy of models. A classification model which is based on machine learning classification matrix to assess the degradation of bearing is proposed to improve the accuracy of classification model. Review work demonstrates the comparisons among the available state-of-the-art methods. In the end, unexplored research technical challenges and niches of opportunity for future researchers are discussed.

Keywords

degradation states, health condition indicator, machine learning, diagnostic model, prognostic model.

1. Introduction

Machine learning classification matrix is used to evaluate the degradation states of bearing over time due to variation in operating parameters and environmental conditions such as load, speed, high temperature, etc. Data driven machine learning model permits retrieving useful information from rotor bearing system using smart monitoring, multi-feature fusion, health condition indicator and advanced diagnostic and prognostic algorithms. Bearing condition monitoring is important for Industry 4.0 to reduce the economic loss and unscheduled downtime of mechanical systems caused by unexpected failures of bearing. Industry 4.0 is a paradigm shift in the modernization of industry propelled by the ever-growing computational capabilities, technological improvement, accuracy of prediction and recent advances in data driven model.

Diagnostics study the fault detection, fault isolation and fault identification in monitored mechanical machinery whereas prognostics deals with the prediction of fault before it occur. The fault detection is to observe the wrong functioning of machinery, where the fault isolation is to identify the components where fault takes place in complex system.

The fault identification is to indicate the nature of fault whereas the prediction of fault is to determine the evolution of fault in machinery before it reaches a critical stage. It has been observed that with the advancement of software technology, artificial intelligence methods are replacing the traditional diagnostics and prognostics systems to enhance the performance of health monitoring.

This paper has reviewed the publications from the science and engineering journals on bearing diagnostic and prognostic in the past 20 years. The published articles were retrieved mainly from Google Scholar using the search terms "bearing diagnostics and prognostics" and "bearing condition monitoring" and filtered by year, access, citations and relevancy. It is observed that there is an increasing trend in number of publications in this area of research after 2014. There are also some highly cited review papers from researchers at universities and industry experts in the past. A brief review of some of key papers is provided in chronological order. In the past few decades, development in diagnostics and prognostic of industrial systems had been reviewed. The multiple sensor and data fusion techniques used in condition based maintenance decision making [19]. Design methodology had been explored for converting data into prognostic information of rotary machinery system [28]. Prognostic techniques for non-stationary and non-linear rotating systems had been studied. The challenges in implementing prognostics technique in industrial system was discussed [23].

In recent years, some researchers started research in the domain of machine learning algorithm for diagnostic and prognostic. A review on spectral kurtosis theories namely; spectral kurtosis, kurtogram, and

(*) Corresponding author.
E-mail addresses: S. Kumar - satishalwary@gmail.com, P. Kumar - paraskum007@rediffmail.com, G. Kumar - girish.kumar154@gmail.com

protrugram applicable for fault identification in bearing had been presented [52].Various aspects from data acquisition to remaining useful life prediction (RUL) in the field of machinery prognostic had been discussed. Authors divided the machinery prognostic program into various stages namely; data acquisition, development of health indicator, division of health stages [29].Generalized gamma distribution was used for the prediction of corrective maintenance of fleet vehicle [5]. Deep learning applications for system health management had demonstrated the benefits of deep learning for fault diagnosis and prognosis is [24]. Bearing diagnostics approaches were compared to consider the impulse behavior of vibration signal. The first approach was considered preprocessing the probabilistic component of the vibration signal by employing the minimum entropy de-convolution approach and the spectral kurtosis method. The second approach was considered the modeling of cyclo-stationarity based on spectral coherence and spectral frequency [1]. A comprehensive survey on recent development in vibration data fusion and application of deep learning tools in machinery prognosis and discussed the identification of research trend, unexplored challenges were provided[10].This paper gives a comprehensive review on bearing diagnostics and prognostics.

Hence, authors attempted to summarize a review for targeted journals published from 2000-2019. The literature search has been done among the electronic database i.e. Science Direct, IEEE explorer and Scopus. The published journals had been explored in the search engine with the following words: diagnosis, prognosis and condition monitoring, degradation model.

The main contributions of this paper are as following:

- The paper reviews the different health condition indicators used in the prior state of art for degradation assessment of bearing.
- The paper also reviews the diagnostic and prognostic models for the remaining useful life estimation of the bearing.
- The paper also discusses the case study of classification model to improve the accuracy of degradation assessment.

The remaining paper is organized as follows. In section 2, data acquisition and health condition indicator is explored for degradation assessment of bearing. Section 3 discusses the diagnostic models. Section 4 focuses on prognostic models for RUL prediction. Section 5 discusses the case study of classification model. Finally, Section 6 concludes the research challenge and provides directions for future research trends in the area of diagnostics and prognostics.

2. Data acquisition

Data acquisition is a procedure of acquiring and storing useful data from different sensors mounted on the machinery. Accelerometer sensor is installed to acquire monitoring data which reflects the degradation stages of bearing. Industry 4.0 has started using advanced sensor to capture monitored data for accurate maintenance decisions. Following obstacles i.e. interferences from operating conditions, noise, cost, time, service period and unexpected failure are main factors for decrease in the quality of the data. Data sources will be helpful to researchers to develop data driven diagnostic & prognostic models.

2.1. Experimental data

Experimental prognostic data set is acquired from accelerated degradation test. This paper has selected bearing prognostic datasets from the repository of NASA, Franche-Comte Electronics Thermal Science and Optics-Sciences and Technologies (FEMTO) and Case Western Reserve University (CWRU) bearing dataset.

Intelligent Maintenance system (IMS) bearing dataset**:** Bearing degradation data was generated by the center for intelligent maintenance system, university of Cincinnati with support from Rexnord corp. (www.imscenter.net). Four (Rexnord ZA-2115 double row) bearings were used in the experiment. All bearings used in the experiment are lubricated. Two accelerometers (PCB353B33 High sensitivity quartz ICP) are installed on the bearing housing to collect the horizontal and vertical vibration signals generated from the bearing respectively. Three run-to-failure tests are conducted to generate three data sets in different time periods. The test 2 consists of 984 files generated by recording data at every 10 minute with the help of NI DAQ card 6062E. The experiment is stopped when a significant amount of metal debris is found on the magnetic plug of the tested bearing. In this paper, test 2 data is utilized for the analysis of bearing degradation condition.

FEMTO bearing dataset**:** This dataset had been provided by FEMTO and was shared in the IEEE international conference (www.femto-st.fr/f/d/IEEEPHM2012- Challenge-Details.pdf). The data is collected from 17 run-to-failure data of rolling element bearing captured from accelerated degradation test in few hours. An accelerometer and a thermocouple were employed to acquire the vibration signals and the temperatures respectively. The healthy bearing was allowed to degrade naturally without introducing a fault. During experimentation, the frequency resolution and time length of each sample were maintained at 10 Hz and 0.1 s respectively. The bearing useful life is estimated at a threshold when vibration signal exceeds 20g.

Case Western Reserve University (CWRU): This data set had been provided publicly from a CWRU. In this data set, electro-discharge machining was used to create an artificial fault in deep groove roller bearing with fault depth of 0.1778 mm, 0.3556 mm, and 0.5334 mm. They acquired vibration data at a sampling frequency 12 kHz and each data sample containing 2048 points. This data set has been used in fault identification, signal processing and machine learning for bearing fault detection. The properties of experiments form two historical data sources are provided in Table 1.

Table 1. Properties- Experimental data

Data	Bearing	Rotor diameter	Load	Speed	Sampling frequency	Reference
IMS	Rexnord ZA-2115	0.331 inches	6000 lb	2000 RPM	20 kHz	www.imscenter.net
CWRU	6205-2RS	0.007 inches	0-3 hp	1797r/min	12kHz	http://csegroups.case.edu

The data source, parameter properties of the bearing has been provided to facilitate researchers for the development of degradation model of the bearing from experimental data.

2.2. Health condition indicator

Health condition indicator give information about the state of the machinery by analyzing the information contained in signal. This indicator is expected to quantify the degradation of machinery. A suitable health condition indicator is to follow the degradation trend for precise prediction results. This section provides the detailed information on the development of health condition indicator used in the evaluation of bearing performance.

Commonly, a health condition indicator is developed physically and virtual to serve as a quantitative description of bearing health condition. Confidence value was used to identify the current degradation state through self- organizing map[17]. Health assessment indicator had been developed based on negative log likelihood probability that measures bearing performance degradation[59]. Probability of degradation was discussed as health indicator[38]. Probability of different

Table 2. Health Condition Indicator

Health Indicator	Technique	Reference
Confidence Value	Self organising map	[17]
Health assessment Indication	Gaussian mixture model based negative log likelihood probability	[59]
Probability	Hidden Markov modeling	[38]
Generalized dimensionless bearing health indicator	upper and lower bounds non-central chi-square distribution	[53]
Monitoring Index	Self organising map	[29]
Virtual Health Index Current Tracking metric	Locally weighted linear regression method	[30]
Maximum likelihood ratio	Statistical methodology	[6]
Dimensionless health indicator	Linear rectification	[16]
Predication bandwidth	Multi scale convolutional neural network	[17]
Average spectral radius, the maximum eigen value, number of random point	Random matrix single ring machine learning	[18]
Moving average cross-correlation coefficient power spectral density	Measure the similarity of power spectral density of signal with adjacent signal	[19]

degradation states were calculated through Hidden Markov modeling. Dimensionless health indicator was presented to assess current health condition of the bearing [50]. They calculated mathematically upper and lower bound of the dimensionless health indicator by using non-central chi-square distribution. Condition monitoring index based on self-organizing map was developed to detect incipient bearing faults quickly [29].Virtual health index was introduced based on locally weighted linear regression method [30]. Statistical methodology based on maximum likelihood ratio was presented to design condition indicators [6]. These condition indicators had shown high potential to describe different phases of bearing degradation process. Dimensionless health indicator was developed through linear rectification technique [2]. They defined the indicator as the ratio of root mean square of the vibration acceleration signal at current time to the root mean square of the vibration acceleration signal for the baseline condition. New health condition indicator was developed based on combination of bathtub curve, multi scale convolutional neural network and inverse hyperbolic tangent function [55]. Three degradation indicators were proposed i.e. average spectral radius, the maximum eigen value and number of random points in inner ring by random matrix single ring machine learning. A data source matrix had been constructed from roller bearing full life failure experimental data through normalizing, singular value decomposition [37]. Health state was evaluated with moving average cross-correlation coefficient power spectral density of signal [56]. A new health index had been proposed on moving average cross-correlation of energy distribution of a signal in the frequency domain. It can also distinguished different health state by determining failure threshold for each case. Table 2 summarizes the health condition indicator (s) used in the published journals such as Mechanical system and signal processing, Journal of sound and vibration, Maintenance and reliability, Computer and Industrial engineering, Reliability Engineering and system safety, Microelectronics reliability and Digital Signal Processing A Review Journal.

This section focuses on the data acquisition and health condition indicators of the bearing vibration data. Then, different health condition indicators are described. This study will be helpful for researchers in the bearing diagnostic and prognostics.

3. Diagnostic model

Diagnostic model is vital to diagnose faults in machineries quickly and precisely. The model learns from the selected features obtained from the health condition of machinery. These models can be categorized in two class i.e. classification and regression based. There is a paradigm shift in research trends from traditional diagnostics approach to machine learning based diagnostics approach. This section investigates recent classification and regression based diagnostic models used in the health condition monitoring of machinery.

3.1. Classification and Regression model

The classification model gives an output in term of categorical variable using class labeled input data obtained from the fault feature extracted from machinery. Regression approach is used to explain the relationship between one continuous dependent variable and multiple independent variables. Regression model gives an output in terms of numeric variable. Model parameters are iteratively updated through optimization algorithm and classes are equally distributed to avoid bias and generalization capability. There are various intelligent classification model i.e. artificial neural network, support vector machine, fuzzy sets theory; fuzzy set theory based expert system used in the diagnosis of the bearing health condition. Artificial neural network on vibration signal data was used for bearing fault diagnosis[20].They trained the model to classify seven different bearing classes and classifier produce the high accuracy diagnosis of real bearing defects. Artificial neural network was discussed for fault diagnosis of rolling element bearing [21]. They trained the model through back propagation algorithm with a subset of experimental data obtained from machine condition. An approach based on dimensional exponent integrated with a surrogate data testing was suggested for bearing condition diagnosis [22].

Novel rough support vector data description method was designed for bearing performance degradation assessment based on one-class classifier [23]. They removed the problems like sensitive to outliers, over-fitting and invariability of model parameters with time. Orbit pattern recognition algorithm using the deep learning proposed for rotating machinery diagnostic [24]. They classified the fault modes of rotating machinery through orbit images. Improved support machine was developed for fault diagnosis based on multi classification of the condition [25]. They proposed an improved voting scheme in one-against-one support vector machine to improve the classification accuracy. Dynamic time warping in machine learning algorithm was proposed to bearing fault classification for mechanical fault detection [26]. They compared the accuracy with the traditional machine learning algorithm. Support matrix machine was developed for roller bearing condition monitoring using matrix as input. However, data was distinguished effectively by two parallel hyper planes and result showed that support matrix machine had better recognition performance as compare to support vector machine [27]. Multi-step support

Table 3. Diagnostic models

Category	Model	Major contributions	Reference
Classification & Regression	Artificial neural network	Back propagation algorithm	[44]
	Computational scheme	Surrogate data testing	[18]
	Rough support vector data	One-class classifier	[65]
	Deep learning	Orbit pattern recognition	[21]
	Improved support machine	Multi classification	[36]
	Machine learning algorithm	Dynamic time warping	[47]
	Support matrix machine	Distinguish the data with two Parallel hyper-planes	[41]
	Multi-step support	Update feature vector	[58]
Hybrid	Principal component analysis & Least square support vector machine	Multi-feature fusion technique	[11]
	Locality preserving projection & Gaussian mixture	Negative log likelihood	[8]
	Kernel principal component analysis, autoregressive & support vector machine	Fault location with recognition rate	[62]
	Weight sparse	Combined the sparse and kurtosis of the envelope spectrum	[61]
	Vibration Resonance	Resonance of main and auxiliary signal	[13]
	Fuzzy C-means and K-nearest neighbor	Using relatively amount of data	[12]
	SOM and PCA	Frequency analysis of residues produced by hybrid algorithm	[27]
	Linear discriminant analysis & Pattern recognition	Two dimensional visualization	[62]

vector regression was proposed method for fault diagnosis in the rotatory machinery [28].

3.2. Hybrid model

Hybrid model combined the advantages of different diagnostic models through their integration. This category contains limited number of publications. Bearing process degradation was studied using the principal component analysis and optimized least square support vector machine[29]. Original features were merged and new features were produced with the help of multi-features fusion technique, principal component analysis. Fault diagnosis had been analyzed by combining locality preserving projection and Gaussian mixture models [30].They developed negative log likelihood probability to quantify the bearing performance gradation using the Gaussian mixture model. Multi-feature fusion diagnosis approach was proposed based on the combination of Kernel principal component analysis, autoregressive model and particle swarm optimized support vector machine [31]. They identified the fault location with recognition rate, generalization capability in small training samples and degree of performance degradation of roller bearing.

Fault diagnosis of rolling element bearing was developed based on linear discriminant analysis and pattern recognition [32]. Two-dimensional visualization and classification accuracy of bearing data were showed to identifying different fault categories effectively. Weight sparse model was presented for bearing fault diagnosis [61]. The coefficient sequence of fault information i.e. sparse and kurtosis of the envelope spectrum were combined to develop weight sparse model. Two model gradient descent and Bayesian were combined to develop a hybrid algorithm. Vibration resonance method was developed on bearing fault diagnosis [13]. The experimental and simulated vibration signals were analyzed on bearing fault diagnosis. Fuzzy C-mean and optimized K-nearest neighbor method was combined to make accurate judgment of bearing fault [12]. The basic features were classified using small amount of fault data. Fault diagnosis of bearing was analyzed two dimensional with the help of linear discriminate analysis and pattern recognition [62].

Hybrid algorithm of SOM and PCA was developed to extract the bearing fault category [27]. The proposed algorithm had isolated the characteristic frequency of bearing fault from residue of data. Table 3 summarizes the different diagnostic models developed in the literature.

This section discusses the diagnostic models used in the bearing diagnosis. Review of the diagnostic models will help to researchers to solve industrial applications.

4. Prognostic model

The prognostic model is used to forecast the remaining useful life (RUL) of machinery before the machinery reach the failure stage based on the health condition information. The RUL of any system is defined as the time duration from present time to the functional failure of machine. Dynamic models were developed for the development of reliable prognostic algorithm [34]. Dynamic models were used to predict the changes in dynamic behavior reflecting the fault type and severity. Prognostic models are classified in to three broad category i.e. statistical data driven prognostic model, hybrid model and machine learning method.

4.1. Statistical data driven model

Statistical data driven prognostic model is based on empirical knowledge to estimate the remaining useful life of machinery. Statistical model is useful to study the uncertainties in the degradation process of machinery and its influence on the prediction of remaining useful life. Present work discusses the recent developments in the various traditional statistical prognostic models like Gaussian Hidden Markov, Gaussian process model, Wiener process model, inverse Gaussian process model and dynamic regression model. Statistically data driven approaches were reviewed for RUL estimation [45]. The pros and cons of the recent model developments and classification of RUL estimation model were discussed. Mixture of Gaussians Hidden Markov model was used for prognostics of bearing [48]. They generated the complex emission probability density function from the wavelet packet coefficients feature extracted from the raw vibration

Table 4. Pros and cons of prognostic model

Type of Prognostic model	Pros	Cons	References
Statistical data driven model	Dynamic calibration of model to adopt to evolving trend	Need for extremely large amount of data in numerous operating conditions Generalization capabilities are undefined	[45] [39] [7][40][49][54][2] [45]
Hybrid model	Identify the nonlinear relationship between the variables	Unable to learn from clustering integration	[48][15][57][49][50][51]
Machine Learning model	Solve the problem of gradient disappearance Construction of sample pair with advanced algorithm Consider the time cumulative effect of historical information on future information from a structural perspective.	Unable to learn long-time timing information	[42][9][25][60][14] [64][58]

signal. The parameters of model were estimated which best fit the degradation phenomenon. Switching Kalman filter approach was used for the prognostic of roller bearing [43]. This approach uses multiple dynamical models each describing a different degradation process.

Likelihood distribution obtained in the Gaussian process following Bayes' rule was used to estimate the RUL of bearing [7]. Inverse Gaussian process model with random effect was discussed to estimate the RUL of bearing [40]. Degradation model parameters were updated by Bayesian method which can capture the real condition of the system. Statistical model was developed for different stages of bearing degradation signal [51]. They considered the drift coefficient at current time in the likelihood function. Monte Carlo simulation was used to develop an RUL prediction approach [54] . Multiple change point wiener process model was developed as a degradation model. Recursively updated dynamic regression model was used to estimate the RUL of bearing [2]. They demonstrated experimentally that excellent prognostic performance of dynamic regression model due to its ability to determine time to start prediction and dynamic calibration of model.

4.2 Hybrid model

Hybrid prognostic model is an attempt to integrate the advantages of different prognostic models. There is limited literature available under this category. Principal component analysis and optimized least squares support vector machine based approach was proposed for bearing degradation prediction [11]. The original features were merged by principal component analysis and optimized the model parameters by particle swarm optimization. Remaining useful life prediction methodology that utilizes mechanistic modeling of vibration and self-training of parameter adaptation was suggested [31]. Prognostic approach utilizing fuzzy adaptive resonance theory map, neural network and Weibull distribution was proposed for RUL prediction [3]. The learned nonlinear time series and seven classes were defined for bearing degradation.

Hybrid prognostic model was developed for health monitoring using bond graph framework [22]. Variance adaption scheme with a statistical model was proposed for system parameter. The effective prediction of the RUL was produced with in confidence bounds. Grey Markov model was used to predict the RUL of roller bearing [35]. The fractal spectrum parameters were used to generate degradation trend and predict the RUL with higher prediction accuracy. Hybrid model of principal component analysis and internet of things with multi sensor was used to predict the bearing life [15]. Multi-dimensional feature predication algorithm had described the life information of rolling bearing from various angles as compare to single time domain feature predication. Hybrid model of support vector machine and degradation tracking model was presented to improve the accuracy of RUL [57]. Features were dimensionless and prognostic works solve the problem of time to start prediction and random fluctuation of measurement.

4.3. Machine Learning model

Machine learning prognostic models study the degradation trend of machinery using machine learning techniques i.e. artificial neural network, support vector machine, web semantic tool, long short-term memory and recurrent neural network. This paper considers only the recent research development in the field of machine learning techniques. Innovative prognostic model based on health state probability estimation was presented [25]. Health state probability was estimated by support vector machine for RUL prediction. Deep learning approach was discussed to predict the RUL of bearing based on deep auto encoder and deep neural networks [42]. They presented deep auto encoder joints features compression to retain effective information without increasing the scale. Recurrent neural network based on encoder-decoder framework with attention mechanism was proposed to predict automatic health indicator which were designed with the RUL values [9]. Features were extracted from five band-pass energy values of frequency spectrum. Proposed method was achieved lowest average percent error and highest average score as compare to traditional method. Accurate RUL prediction was depending on the use of long time-dependent information from the long-time sequence data effectively. Long short-term memory recurrent neural network was used to predict the RUL of bearing [60]. Degradation states were identified by giving input into long short-term memory recurrent network. Bearing performance degradation was studied using long short-term memory with multi-resolution singular value decomposition (MRSVD). The decomposition of vibration signal with MRSVD and reconstruction help to accurately identify the fault point in vibration signal [14]. New data driven transfer learning RUL prediction approach was proposed to solve the distribution discrepancy problem [64]. The fault occurrence time was detected by hidden Markov model. The domain discrepancy metrices and domain classifier were used to acquire domain invariant features through domain adaption module and condition recognition. New approach was presented to predict the RUL of industrial roller bearing based on state recognition and similarity analysis [16]. Life proportional adjustment function was constructed with the help of comprehensive similarity analysis between historical bearing data and monitoring bearing data. Life model was constructed by defining state matrix of different operation states of roller bearing. Result showed that proposed approach had better prediction accuracy and generalization as compare to hidden Markov model and grey model. Predicated fatigue life of radial cylindrical roller bearing subjected to radial and axial load was discussed [59]. Remaining useful life prediction was made via the combined use of support vector machine as a classification tool and autoregressive integrated moving average based identification. An expert tool was used for real time monitoring to prevent the potential failure of machines [26]. A novel model combines the importance of machine criticality assessment criteria with interaction between them was proposed [20]. Remaining useful life of a ball bearing was predicated using classification and regression techniques. Machine learning principles was

Table 5. Prognostic Models

Category of Prognostic model	Comments	Reference
Statistical data driven	Statistical data driven approach used for classification of degradation state	[45]
	Switching Kalman filter approach employed to identify unstable degradation state	[43]
	• Gaussian process applied to estimate posterior distribution of the bearing relative time • Probability density function was calculated for posterior distribution using Bayes rule • Gaussian process model used to evaluate likelihood	[7]
	• Inverse Gaussian model was employed with random effect to characterize the degradation process of the system • Parameters were updated by Bayesian method	[40]
	Statistical model was used to find analytical expression for posterior drift distribution	[49]
	• Monte Carlo simulation algorithm considered for multiple change-point • Wiener process employed to construct degradation model • Bayesian approach applied for parameters estimation • Exact recursive model was used for updation	[54]
	• Dynamic regression model used to forecast start time • Predication were made using alarm bound technique • Predicating future health indicator values by recursive updation • Remaining useful life estimation using time steps to fail threshold	[2][45]
Hybrid model	Mixture of Gaussians Hidden Markov models were used for better implementation and interpretability	[48]
	Hybrid model include internet of things with multi sensors was used for PCA	[15]
	Hybrid degradation tracking model (support vector machine and hybrid degradationtracking model)	[57]
Machine Learning	• Deep learning use of subset based deep auto encoder • Feature compression	[42]
	Novel deep learning use recurrent neural network based on Encoder-decoder framework with attention mechanism	[9]
	• Health state probability estimation using support vector machine • Prognostic model parameters were updated using historical knowledge	[25]
	Long short memory Recurrent neural network proposed for automatic detection and to identify fault occurrence	[60]
	Long short memory network with multi-resolution singular value decomposition technique used to detect accurate fault in vibration signals	[14]
	Training and testing data distribution discrepancy problem was solved by Transfer Learning based on multiple layer perceptron	[64]
	State recognition and similarity analysis used clustering algorithm and threshold correction to solve the problem of prediction accuracy and generalization	[16]

used to develop an algorithm to recognize underlying mapping function [46]. Table 4 summarizes the pros and cons of prognostic model. Table 5 summarizes the main features of prognostic models.

This section discusses the different prognostic models used in the bearing prognostics. The pros and cons of prognostic models are also discussed.

5. Case Study

Roller bearings have been degraded from normal condition to failure condition with the duration of time due to harsh industrial working conditions. However, rolling bearings have a low ability to withstand impact, so their service life is uncertain. This paper explores multistage categorization of bearing degradation. Table 6 summarizes the stages for bearing degradation criteria. Bearing degradation is categorized in one of the following three stages. The degradation states of bearing over time is shown in Fig.1.

- Stage I : Healthy stage
- Stage II : Degradation stage
- Stage III : Critical stage

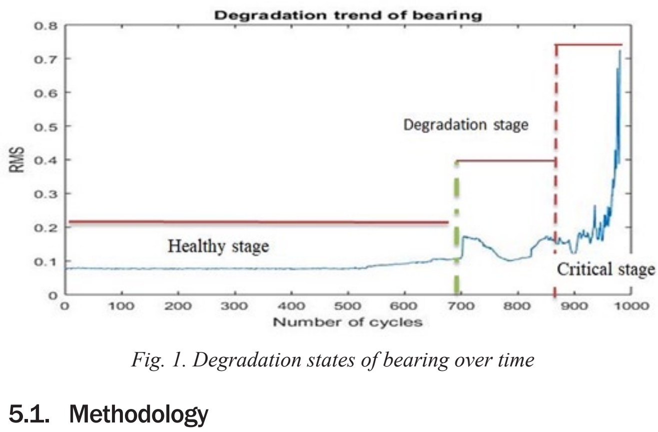


Fig. 1. Degradation states of bearing over time

5.1. Methodology

The health data used for bearing degradation assessment is the bearing vibration signal. Firstly, statistical features are extracted from the vibration signal data and features are selected from the values of correlation coefficient. Secondly collect the samples and divide the samples into the training samples and testing samples. Then the testing data used for input into a classification model for bearing degra-

Table 6. Stage for bearing degradation criteria

Stage	Criteria: Severity of impact in the bearing degradation
Catastrophic	Any impact which could potentially cause the loss of primary system functions resulting in significant damage to the mechanical system and cause the loss of life
Critical	Any impact which could potentially cause the loss of primary system functions resulting in significant damage to the mechanical system and negligible loss to life
Degradation	Any impact which degrades system performance functions without appreciable damage to either mechanical system or life
Healthy	Any event which could not cause degradation of system performance function(s) resulting in negligible damage to mechanical system

Fig. 2. Flowchart of proposed methodology

dation assessment. The flow chart of the bearing classification model proposed is shown in Fig. 2.

5.2. Experimental data

The bearing degradation data was generated by the center for intelligent maintenance system (IMS), university of Cincinnati with support from Rexnord corp. The used data set in this paper is downloaded from prognostics center of excellence through prognostic data repository. The bearing used in the experiment is Rexnord ZA-2115 double row bearing to support a rotating shaft. The bearing test rig and accelerometer sensor placement are shown in Fig. 3.

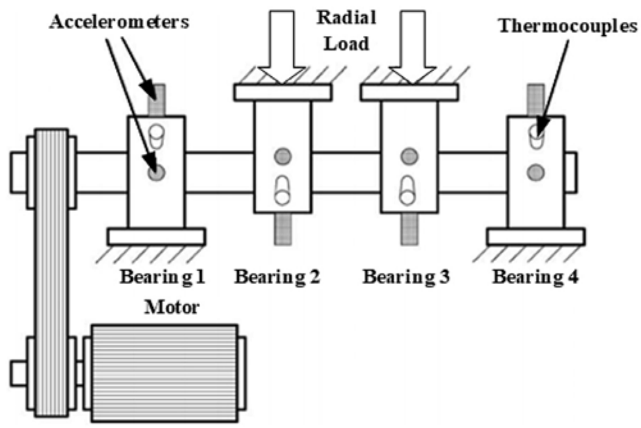


Fig. 3. Bearing Test rig [3]

The bearing test rig was designed to generate run-to-failure data from Feb.12, 2004 to Feb.19, 2004. Four bearings are used in this experiment. The rotating speed of the shaft is kept constant at 2000 rpm with the help of alternative current motor coupled to the shaft via rubber belts. The radial load of 6000 lb is applied onto the bearings by a spring mechanism. All bearings used in the experiment are lubricated. Two accelerometers (PCB353B33 High sensitivity quartz ICP) are installed on the bearing housing to collect the horizontal and vertical vibration signals generated from the bearing respectively. The data sampling frequency is 20 kHz. Three run-to-failure tests are conducted to generate three data sets in different time periods. The test 2 consists of 984 files that are 1-s vibration signal snapshots recorded at every 10 minute with the help of NI DAQ card 6062E. Each file stored of 20480 points with the sampling frequency set at 20 kHz. The failure in the bearing had occurred when the bearing cross the designed life time of the bearing which is more than 100 million revolutions. The experiment is stopped when a significant amount of metal debris is found on the magnetic plug of the tested bearing. In this paper, test 2 data are used to the time domain feature to monitor the bearing degradation condition.

5.3. Classification model

Classification based models have been used to develop a relationship between independent variables (i.e. features vectors) and dependent variables (i.e. response in term of predefined stages identified by labels). In this paper, we have feed the input feature vectors values (Predicators) and response corresponding with the bearing degradation system stages into classification learner tool in Matlab to obtain predicated label for future time period. The input feature vectors values are extracted from preprocessed time series of bearing vibration signals relevant to the degradation stages of the bearing. We have categorized the degradation of the bearing into three classes such as healthy stage, degradation stage and critical stage. This paper has considered 15 predicators such as kurtosis, peak, crest factor, standard deviation, variance, rms, mean, mode, median, range of values, mean and median absolute deviation, peak amplitude to rms ratio, interquartile range, root sum of square level and maximum to minimum difference. Three response classes are used to predict an output categorical variable using labeled input data. Predicators are independent to each other. In supervised machine learning methodology, the classification labels have been assigned to the feature vector values to which training instances belong. In this paper, we have found the accuracy of models in different classification algorithms. Table 7 shows the accuracy results of different classification algorithms.

Table 7. Comparison results on the testing accuracy

Classifier	Accuracy (%)
Decision tree	96.1 %
Linear discriminant analysis	92.9 %
SVM (Cubic SVM)	96.5 %
SVM (Medium Gaussian SVM)	96.2 %
Nearest neighbor classifiers (Weighted KNN)	93.5%
Ensemble classifiers (Bagged tree)	96.5 %

This section discusses the different classification models used for bearing health analysis. The accuracy of different classification algorithms are also discussed to study the degradation stages in bearing.

5.4. Bearing degradation failure

In the case study, the bearing degradation failure is discussed with machine learning classification matrix. Vertical axis and horizontal axis denote the actual and predicted label respectively. Elements in the main diagonal are the classification accuracies and others are the classification errors. Bearing degradations are divided into four categories i.e. healthy, degradation, critical and catastrophic. For each conditions, there are 1-second snapshots, each of which consists of

20,480 points. Each samples has 2000 data points. Totally, there are 1400 samples for the four health conditions. Random 50 % samples are for training and the remaining samples are for testing.

Graph theoretic model was considered the system structure explicitly and applied to model functions using matrix approach to examine the cause and effect. Result showed the reliability enhancement using step-by-step methodology [32]. Structural graph model for reliability at various hierarchical levels was developed by converting reliability graph into equivalent matrix [33]. System model was developed incorporating four states of degradation for each component [39].

6. Technical challenge

It has been observed that despite the positive outcomes from the existing state of art, there are many current research challenges need to be addressed. More research should be conducted on incorporation of uncertainty in diagnostic and prognostic model. The key idea of data acquisition is to transfer the knowledge gained from experimental data to improve the accuracy of predication model used in industry. It is emphasized to develop information fusion from multi-dimension data. The research problems pertinent in this field are design of health condition indicator and real time model to study the real time degradation of bearing. The big data provides research challenges to build robust diagnostic and prognostic model from machine learning technology. Further, research is required for dimensionless health indicator which is more sensitive to an incipient bearing defect. There is a paradigm shift in research direction from constant operating conditions to variable operating conditions. Thus, it is important to analyze uncertainties caused by time varying operation conditions. To address this change, future researchers need to redefine the failure threshold limit, health state division, degradation model and quantification of uncertainty according to the variable load, speed, etc. Research is needed to determine failure threshold limit for virtual dimensionless health indicator. This section discusses the unexplored technical challenges existing in the bearing health monitoring to match with variable operating condition and advanced machine learning algorithm.

7. Conclusions

In this paper, the prior state of art in the field of diagnosis and prognosis of bearing with emphasis on machine learning based techniques has been summarized. The review is focused on data acquisition, health condition indicator, diagnostic models and prognostic models. The advantage and disadvantage of models, algorithms are presented in this study. A case study is discussed based on machine learning classification matrix to improve the accuracy of degradation assessment. The future research challenges are moving from tradition algorithm to advanced machine learning algorithm to build accurate and robust prediction model. Further, there is a need to develop virtual dimensionless health indicator, failure threshold limit which can match the degradation trend of the bearing.

References

1. Abboud D, Elbadaoui M, Smith W A, Randall R B. Advanced bearing diagnostics: A comparative study of two powerful approaches. Mechanical Systems and Signal Processing 2019; 114: 604-627, https://doi.org/10.1016/j.ymssp.2018.05.011.
2. Ahmad W, Khan S A, Islam M M M, Kim J M. A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models. Reliability Engineering and System Safety 2019; 184: 67-76, https://doi.org/10.1016/j.ress.2018.02.003.
3. Ben Ali J, Chebel-Morello B, Saidi L et al. Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. Mechanical Systems and Signal Processing 2015; 56: 150-172, https://doi.org/10.1016/j.ymssp.2014.10.014.
4. Ben Ali J, Fnaiech N, Saidi L et al. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. Applied Acoustics 2015; 89: 16-27, https://doi.org/10.1016/j.apacoust.2014.08.016.
5. Andrzejczak K, Selech J. Generalised gamma distribution in the corrective maintenance prediction of homogeneous vehicles. Lecture Notes in Networks and Systems 2019; 68: 519-529, https://doi.org/10.1007/978-3-030-12450-2_50.
6. Antoni J, Borghesani P. A statistical methodology for the design of condition indicators. Mechanical Systems and Signal Processing 2019; 114: 290-327, https://doi.org/10.1016/j.ymssp.2018.05.012.
7. Boškoski P, Gašperin M, Petelin D, Juričić D. Bearing fault prognostics using Rényi entropy based features and Gaussian process models. Mechanical Systems and Signal Processing 2015; 52-53(1): 327-337, https://doi.org/10.1016/j.ymssp.2014.07.011.
8. Chen L, Hang Y. Performance degradation assessment and fault diagnosis of bearing based on EMD and PCA-SOM. Vibroengineering Procedia 2013; 2: 12-16.
9. Chen Y, Peng G, Zhu Z, Li S. A novel deep learning method based on attention mechanism for bearing remaining useful life prediction. Applied Soft Computing Journal 2020; 86: 105919, https://doi.org/10.1016/j.asoc.2019.105919.
10. Diez-Olivan A, Del Ser J, Galar D, Sierra B. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. Information Fusion 2019; 50: 92-111, https://doi.org/10.1016/j.inffus.2018.10.005.
11. Dong S, Luo T. Bearing degradation process prediction based on the PCA and optimized LS-SVM model. Measurement: Journal of the International Measurement Confederation 2013; 46(9): 3143-3152, https://doi.org/10.1016/j.measurement.2013.06.038.
12. Dong S, Xu X, Chen R. Application of fuzzy C - means method and classification model of optimized K - nearest neighbor for fault diagnosis of bearing. Journal of the Brazilian Society of Mechanical Sciences and Engineering 2016; 38(8): 2255-2263, https://doi.org/10.1007/s40430-015-0455-9.
13. Gao J, Yang J, Huang D et al. Experimental application of vibrational resonance on bearing fault diagnosis. Journal of the Brazilian Society of Mechanical Sciences and Engineering 2019; 41(1): 1-13, https://doi.org/10.1007/s40430-018-1502-0.
14. He M, Zhou Y, Li Y et al. Long short-term memory network with multi-resolution singular value decomposition for prediction of bearing performance degradation. Measurement 2020; 156: 107582, https://doi.org/10.1016/j.measurement.2020.107582.
15. Heng W, Guangxian N, Jinhai C, Jiangming Q. Research on Rolling Bearing State Health Monitoring and Life Prediction Based on PCA and Internet of Things with Multi-sensor. Measurement 2020: 107657, https://doi.org/10.1016/j.measurement.2020.107657.
16. Heng W, Jinhai C, Jiangming Q, Guangxian N. A new approach for safety life prediction of industrial rolling bearing based on state recognition and similarity analysis. Safety Science 2020; 122:104530, https://doi.org/10.1016/j.ssci.2019.104530.
17. Hong S, Zhou Z, Zio E, Hong K. Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method. Digital Signal Processing: A Review Journal 2014; 27(1): 159-166, https://doi.org/10.1016/j.dsp.2013.12.010.
18. Janjarasjitt S, Ocak H, Loparo K A. Bearing condition diagnosis and prognosis using applied nonlinear dynamical analysis of machine vibration signal. Journal of Sound and Vibration 2008; 317(1-2): 112-126, https://doi.org/10.1016/j.jsv.2008.02.051.
19. Jardine A K S, Lin D, Banjevic D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing 2006; 20(7): 1483-1510, https://doi.org/10.1016/j.ymssp.2005.09.012.

20. Jasiulewicz-Kaczmarek M, Antosz K, Żywica P et al. Framework of machine criticality assessment with criteria interactions. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 23(2): 207-220, https://doi.org/10.17531/ein.2021.2.1.
21. Jeong H, Park S, Woo S, Lee S. Rotating Machinery Diagnostics Using Deep Learning on Orbit Plot Images. Procedia Manufacturing 2016; 5: 1107-1118, https://doi.org/10.1016/j.promfg.2016.08.083.
22. Jha M S, Dauphin-Tanguy G, Ould-Bouamama B. Particle filter based hybrid prognostics for health monitoring of uncertain systems in bond graph framework. Mechanical Systems and Signal Processing 2016; 75: 301-329, https://doi.org/10.1016/j.ymssp.2016.01.010.
23. Kan M S, Tan A C C, Mathew J. A review on prognostic techniques for non-stationary and non-linear rotating systems. Mechanical Systems and Signal Processing 2015; 62: 1-20, https://doi.org/10.1016/j.ymssp.2015.02.016.
24. Khan S, Yairi T. A review on the application of deep learning in system health management. Mechanical Systems and Signal Processing 2018; 107: 241-265, https://doi.org/10.1016/j.ymssp.2017.11.024.
25. Kim H E, Tan A C C, Mathew J, Choi B K. Bearing fault prognosis based on health state probability estimation. Expert Systems with Applications 2012; 39(5): 5200-5213, https://doi.org/10.1016/j.eswa.2011.11.019.
26. Kozłowski E, Mazurkiewicz D, Żabiński T et al. Machining sensor data management for operation-level predictive model. Expert Systems with Applications 2020; 159: 113600, https://doi.org/10.1016/j.eswa.2020.113600.
27. Lamine M, Abdelkrim F. Hybrid SOM - PCA method for modeling bearing faults detection and diagnosis. Journal of the Brazilian Society of Mechanical Sciences and Engineering 2018; 40(5): 1-8, https://doi.org/10.1007/s40430-018-1184-7.
28. Lee J, Wu F, Zhao W et al. Prognostics and health management design for rotary machinery systems - reviews, methodology and applications. Mechanical Systems and Signal Processing 2014; 42(1-2): 314-334, https://doi.org/10.1016/j.ymssp.2013.06.004.
29. Li Z, Fang H, Huang M et al. Data-driven bearing fault identification using improved hidden Markov model and self-organizing map. Computers and Industrial Engineering 2018; 116: 37-46, https://doi.org/10.1016/j.cie.2017.12.002.
30. Li Z, Wu D, Hu C, Terpenny J. An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction. Reliability Engineering and System Safety 2019; 184: 110-122, https://doi.org/10.1016/j.ress.2017.12.016.
31. Liang S Y, Li Y, Billington S A et al. Adaptive prognostics for rotary machineries. Procedia Engineering 2014; 86: 852-857, https://doi.org/10.1016/j.proeng.2014.11.106.
32. Loganathan M K, Gandhi O P. Reliability enhancement of manufacturing systems through functions. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 2017; 231(10): 1850-1868, https://doi.org/10.1177/0954405415612324.
33. Loganathan M K, Gandhi O P. Reliability evaluation and analysis of CNC cam shaft grinding machine. Journal of Engineering, Design and Technology 2015; 13(1): 37-73, https://doi.org/10.1108/JEDT-10-2012-0042.
34. Madar E, Klein R, Bortman J. Contribution of dynamic modeling to prognostics of rotating machinery. Mechanical Systems and Signal Processing 2019; 123: 496-512, https://doi.org/10.1016/j.ymssp.2019.01.003.
35. Meng Z, Li J, Yin N, Pan Z. Remaining useful life prediction of rolling bearing using fractal theory. Measurement 2020; 156: 107572, https://doi.org/10.1016/j.measurement.2020.107572.
36. Miao Q, Zhang X, Liu Z, Zhang H. Condition multi-classification and evaluation of system degradation process using an improved support vector machine. Microelectronics Reliability 2017; 75: 223-232, https://doi.org/10.1016/j.microrel.2017.03.020.
37. Ni G, Chen J, Wang H. Degradation assessment of rolling bearing towards safety based on random matrix single ring machine learning. Safety Science 2019; 118(5): 403-408, https://doi.org/10.1016/j.ssci.2019.05.010.
38. Ocak H, Loparo K A, Discenzo F M. Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics. Journal of Sound and Vibration 2007; 302(4-5): 951-961, https://doi.org/10.1016/j.jsv.2007.01.001.
39. P. Varghese J, Kumar G. Availability Analysis with Opportunistic Maintenance of a Two Component Deteriorating System. International Journal of Materials, Mechanics and Manufacturing 2014; 2(2): 155-160, https://doi.org/10.7763/IJMMM.2014.V2.119.
40. Pan D, Liu J B, Cao J. Remaining useful life estimation using an inverse Gaussian degradation model. Neurocomputing 2016; 185: 64-72, https://doi.org/10.1016/j.neucom.2015.12.041.
41. Pan H, Yang Y, Zheng J et al. Symplectic interactive support matrix machine and its application in roller bearing condition monitoring. Neurocomputing 2020; 398: 1-10, https://doi.org/10.1016/j.neucom.2020.01.074.
42. Ren L, Sun Y, Cui J, Zhang L. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. Journal of Manufacturing Systems 2018; 48: 71-77, https://doi.org/10.1016/j.jmsy.2018.04.008.
43. Reuben L C K, Mba D. Diagnostics and prognostics using switching Kalman filters. Structural Health Monitoring 2014; 13(3): 296-306, https://doi.org/10.1177/1475921714522844.
44. Samanta B, Al-Balushi K R. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. Mechanical Systems and Signal Processing 2003; 17(2): 317-328, https://doi.org/10.1006/mssp.2001.1462.
45. Si X S, Wang W, Hu C H, Zhou D H. Remaining useful life estimation - a review on the statistical data driven approaches. European Journal of Operational Research 2011; 213(1): 1-14, https://doi.org/10.1016/j.ejor.2010.11.018.
46. Singh S, Agarwal T, Yadav G K O M. Predicting the remaining useful life of ball bearing under dynamic loading using supervised learning. Proceedings of the 2019 IEEE, 2019; 1119-1123, https://doi.org/10.1109/IEEM44572.2019.8978649.
47. Sobie C, Freitas C, Nicolai M. Simulation-driven machine learning: Bearing fault classification. Mechanical Systems and Signal Processing 2018; 99: 403-419, https://doi.org/10.1016/j.ymssp.2017.06.025.
48. Tobon-Mejia D A, Medjaher K, Zerhouni N, Tripot G. A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models. IEEE Transactions on Reliability 2012; 61(2): 491-503, https://doi.org/10.1109/TR.2012.2194177.
49. Wang B, Hu X, Li H. Rolling bearing performance degradation condition recognition based on mathematical morphological fractal dimension and fuzzy C-means. Measurement: Journal of the International Measurement Confederation 2017; 109: 1-8, https://doi.org/10.1016/j.measurement.2017.05.033.
50. Wang D, Tsui K L. Theoretical investigation of the upper and lower bounds of a generalized dimensionless bearing health indicator. Mechanical Systems and Signal Processing 2018; 98: 890-901, https://doi.org/10.1016/j.ymssp.2017.05.040.
51. Wang D, Tsui K L. Statistical modeling of bearing degradation signals. IEEE Transactions on Reliability 2017; 66(4): 1331-1344, https://doi.org/10.1109/TR.2017.2739126.
52. Wang Y, Xiang J, Markert R, Liang M. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: a review with applications. Mechanical Systems and Signal Processing 2016; 66-67: 679-698, https://doi.org/10.1016/j.ymssp.2015.04.039.

53. Wang Z, Wu Q, Zhang X et al. A generalized degradation model based on Gaussian process. Microelectronics Reliability 2018; 85(4): 207-214, https://doi.org/10.1016/j.microrel.2018.05.001.
54. Wen Y, Wu J, Das D, Tseng T L. Degradation modeling and RUL prediction using Wiener process subject to multiple change points and unit heterogeneity. Reliability Engineering and System Safety 2018; 176(11): 113-124. https://doi.org/10.1016/j.ress.2018.04.005.
55. Wu C, Feng F, Wu S et al. A method for constructing rolling bearing lifetime health indicator based on multi - scale convolutional neural networks. Journal of the Brazilian Society of Mechanical Sciences and Engineering 2019; 41(11): 1-11, https://doi.org/10.1007/s40430-019-2010-6.
56. Xu L, Pennacchi P, Chatterton S. A new method for the estimation of bearing health state and remaining useful life based on the moving average cross-correlation of power spectral density. Mechanical Systems and Signal Processing 2020; 139: 106617, https://doi.org/10.1016/j.ymssp.2020.106617.
57. Yan M, Wang X, Wang B et al. Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model. ISA Transactions 2019; 98: 471-48, https://doi.org/10.1016/j.isatra.2019.08.058.
58. Yaqub M F, Gondal I, Kamruzzaman J. Multi-step support vector regression and optimally parameterized wavelet packet transform for machine residual life prediction. Journal of Vibration and Control 2013; 19(7): 963-974, https://doi.org/10.1177/1077546311435349.
59. Yu J. Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. Mechanical Systems and Signal Processing 2011; 25(7): 2573-2588, https://doi.org/10.1016/j.ymssp.2011.02.006.
60. Zhang B, Zhang S, Li W. Bearing performance degradation assessment using long short-term memory recurrent network. Computers in Industry 2019; 106: 14-29, https://doi.org/10.1016/j.compind.2018.12.016.
61. Zhang H, Chen X, Du Z, Yan R. Kurtosis based weighted sparse model with convex optimization technique for bearing fault diagnosis. Mechanical Systems and Signal Processing 2016; 80: 349-376, https://doi.org/10.1016/j.ymssp.2016.04.033.
62. Zhang Y, Zuo H, Bai F. Classification of fault location and performance degradation of a roller bearing. Measurement 2013; 46(3): 1178-1189, https://doi.org/10.1016/j.measurement.2012.11.025.
63. Zhao M, Jin X, Zhang Z, Li B. Fault diagnosis of rolling element bearings via discriminative subspace learning: visualization and classification. Expert Systems with Applications 2014; 41(7): 3391-3401, https://doi.org/10.1016/j.eswa.2013.11.026.
64. Zhu J, Chen N, Shen C. A new data-driven transferable remaining useful life prediction approach for bearing under different working conditions. Mechanical Systems and Signal Processing 2020; 139: 106602, https://doi.org/10.1016/j.ymssp.2019.106602.
65. Zhu X, Zhang Y, Zhu Y. Bearing performance degradation assessment based on the rough support vector data description. Mechanical Systems and Signal Processing 2013; 34(1-2): 203-217, https://doi.org/10.1016/j.ymssp.2012.08.008.