УДК: 629.05
DOI: 10.26467/2079-0619-2025-28-2-35-50

To the analysis of methods and mechanisms of predictive modeling of onboard equipment reliability when solving problems of aircraft maintenance workload planning

B.I. Ogunvoul[1], V.D. Budaev[1], D.O. Sizikov[1], N.V. Gorbakon[1], A.V. Vlasova[1]

[1]*Moscow State Technical University of Civil Aviation, Moscow, Russia*

Abstract: The article deals with a method for aircraft maintenance planning based on advanced mathematical modeling techniques. In the course of the research, a mathematical model for forecasting the failure rate of onboard equipment is developed and tested, designed to solve the problems of optimizing decision-making processes for maintenance on the basis of reliability assessment of aviation equipment. The application of Poisson distribution regression in combination with polynomial features allows to reveal the regularities of equipment failures, which depend on operating conditions and maintenance history. For the study, a synthesized dataset was created to simulate different operational scenarios and equipment degradation process. At the first stage, the data were freed from outliers and errors, then normalized to unify the scale of different variables. Next, the data were categorized according to the operating conditions, after which Poisson distribution regression was applied to predict failures. Finally, an efficient maintenance plan that takes into account the predicted failures has been developed using an optimization algorithm. Validation of the model's predictive capabilities and optimization of the maintenance strategy are performed by comparing with archived data on previously performed work. The analysis of the results revealed the peculiarities of the model operation, namely, the application of least squares regression with single coding demonstrates perfect forecasts, which may indicate the need for model transformation and requires additional verification. At the same time, alternative versions of the methodology revealed more realistic error and correlation limits, which also confirms the reliability of the predictive models. The results of the study show that a combined approach using Poisson distribution regression and polynomial signs can significantly improve the accuracy of forecasts. This method, in particular, has demonstrated its effectiveness in modeling onboard equipment failures, which allows to optimize maintenance processes in order to reduce repair costs. The obtained conclusions confirm the possibility of introducing more accurate proactive methods of maintenance planning, which allows to improve aircraft reliability and reduce the inefficiency of their downtime on the ground.

Key words: reliability assessment, aircraft maintenance planning, Poisson distribution regression, predictive modeling of maintenance processes, stochastic processes modeling, operational efficiency, flight safety standards, data-driven maintenance optimization, statistical methods in reliability engineering.

For citation: Ogunvoul, B.I., Budaev, V.D., Sizikov, D.O., Gorbakon, N.V., Vlasova, A.V. (2025). To the analysis of methods and mechanisms of predictive modeling of onboard equipment reliability when solving problems of aircraft maintenance workload planning. Civil Aviation High Technologies, vol. 28, no. 2, pp. 35–50. DOI: 10.26467/2079-0619-2025-28-2-35-50

К анализу методов и механизмов прогностического моделирования надежности бортового оборудования при решении задач планирования объемов работ по техническому обслуживанию воздушных судов

Б.И. Огунвоул[1], В.Д. Будаев[1], Д.О. Сизиков[1], Н.В. Горбаконь[1], А.В. Власова[1]

[1]*Московский государственный технический университет гражданской авиации, г. Москва, Россия*

Аннотация: В статье рассматривается метод планирования технического обслуживания воздушных судов на основе усовершенствованных методов математического моделирования. В ходе исследования разработана и апробирована математическая модель прогнозирования частоты отказов бортового оборудования, предназначенная для решения задач оптимизации процессов принятия решений по техническому обслуживанию на основе оценки надежности авиационной техники. Применение регрессии распределения Пуассона в сочетании с полиномиальными признаками позволяет выявить закономерности отказов оборудования, которые зависят от условий эксплуатации и предыстории технического обслуживания. Для исследования был создан синтезированный набор данных, моделирующий различные сценарии эксплуатации и процесс деградации оборудования. На первом этапе данные были освобождены от выбросов и ошибок, затем нормализованы для унификации масштабов различных переменных. Далее они были разделены на категории в зависимости от условий эксплуатации, после чего применена регрессия распределения Пуассона для прогнозирования отказов. Наконец, с помощью алгоритма оптимизации был разработан эффективный план технического обслуживания, учитывающий прогнозируемые отказы. Валидация прогностических возможностей модели и оптимизация стратегии технического обслуживания осуществляются путем сопоставления с архивными данными о ранее проведенных работах. Анализ результатов выявил особенности функционирования модели, а именно: применение регрессии методом наименьших квадратов с однократным кодированием демонстрирует идеальные прогнозы, что может свидетельствовать о необходимости преобразования модели и требует дополнительной верификации. В то же время альтернативные варианты методологии позволили выявить более реалистичные пределы погрешности и корреляции, что также подтверждает надежность прогностических моделей. Результаты исследования показывают, что комбинированный подход, использующий регрессию распределения Пуассона и полиномиальные признаки, позволяет значительно повысить точность прогнозов. Этот метод, в частности, продемонстрировал свою эффективность при моделировании отказов бортового оборудования, что позволяет оптимизировать процессы технического обслуживания с целью снижения затрат на ремонт. Полученные выводы подтверждают возможность внедрения более точных упреждающих методов планирования ТО, что дает возможность повысить надежность воздушных судов и снизить неэффективность их простоев на земле.

Ключевые слова: оценка надежности, планирование ТО воздушных судов, регрессия распределения Пуассона, прогностическое моделирование процессов ТО, моделирование стохастических процессов, эксплуатационная эффективность, стандарты безопасности полетов, оптимизация процессов ТО на основе данных, статистические методы проектирования надежности.

Для цитирования: Огунвоул Б.И. К анализу методов и механизмов прогностического моделирования надежности бортового оборудования при решении задач планирования объемов работ по техническому обслуживанию воздушных судов / Б.И. Огунвоул, В.Д. Будаев, Д.О. Сизиков, Н.В. Горбаконь, А.В. Власова // Научный вестник МГТУ ГА. 2025. Т. 28, № 2. С. 35–50. DOI: 10.26467/2079-0619-2025-28-2-35-50

Introduction

The aviation industry imposes high requirements on the reliability and safety of flights due to the potential negative consequences of technical failures. Of particular importance is the reliability of aircraft onboard equipment, which has a direct impact on flight safety. Maintenance planning plays a key role in ensuring optimal performance and reliability of aircraft systems. Modern aviation equipment is becoming increasingly complex, which requires the introduction of new, more advanced methods of maintenance planning [1, 2].

One of the most active scientific directions is mathematical modeling, which allows to perform quantitative analysis of onboard equipment reliability taking into account various factors such as failure statistics, operating conditions and applied maintenance strategies [3–5]. The purpose of this study is to develop a mathematical model that allows estimating the reliability parameters of onboard equipment and facilitating the formation of a maintenance program that meets industry standards for flight safety and operational efficiency [6].

The main objectives of this study are:

1) development of a mathematical modeling method for predicting the reliability of aircraft equipment;

2) application of the accumulated experience for parameterization of the developed model for the purpose of reliable forecasting of failure rate and volumes of required maintenance.

As part of the literature review, the known research results on different approaches to assessing the reliability of on-board aircraft equipment are considered.

The following sections of the article will examine in detail the mathematical methods used, present the modeling results, and analyze the impact of the developed maintenance planning program on flight safety and the reliability of aircraft operation.

The aviation industry has traditionally paid special attention to reliability issues. Numerous studies demonstrate the effectiveness of statistical methods such as Poisson regression and Weibull analysis for failure prediction [7]. These methods, which have proven their effectiveness in various industries, provide a basis for improving maintenance planning in aviation. The introduction of onboard condition monitoring systems [8, 9] has contributed to the accumulation of significant amounts of operational data, which has opened new prospects for the development of predictive maintenance based on data analysis models [10, 11].

Modern advances in mathematical modeling have led to the development and implementation of more sophisticated methods, including stochastic models, for predicting reliability and determining maintenance requirements for complex systems. These models take into account not only mean time between failures data, but also operational parameters affecting component degradation processes.

Nevertheless, the problem of integrating these models into a comprehensive maintenance planning system that meets the specific requirements of flight safety standards remains relevant. This study aims to solve this problem by synthesizing classical statistical methods with modern approaches of mathematical modeling to develop an effective maintenance planning system.

Principles of analysis and modeling

This section presents a description of the mathematical model developed within the framework of this study. The model is based on the Poisson process, a widely used approach to modeling countable data for rare events, which include technical systems failures [12].

The mathematical model of onboard equipment reliability is based on the time dependence R(t). To calculate the failure rate λ(t), the Poisson distribution regression method is used, which is based on statistical analysis of product failures data. This method helps to predict the probability of product failures depending on the operating time and a number of other factors, and is determined by the following formula

$$\lambda(t) = f(t)/R(t) = \\ = \exp\left(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k\right), \quad (1)$$

where λ(t) is the expected failure rate, $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$ are regression coefficients, and $x_1, x_2, \ldots, x_k$ are predictors (e.g., operating hours, operating conditions, etc.).

where f(t) is the probability density function of the time-to-failure distribution.

To calculate λ(t), statistical product failures data are used, and the distribution regression parameters are estimated using the maximum likelihood method.

Maintenance planning is formulated as an optimization problem, the goal of which is to minimize the expected downtime D(t), determined by the following formula

$$D(t) = \int_0^t \left(1 - R(s)\right) ds, \quad (2)$$

where R(s) is the reliability function at time s.

Maintenance activities are scheduled at times that minimize D(t), taking into account operational constraints.

The Poisson distribution regression method was chosen to analyze rare events such as equipment failures. This method is appropriate when events occur rarely and can be described in terms of a counting process. The advantage of the Poisson distribution regression is its ability to accurately model the relationship between failure rates and flight hours [12].

The dependent variable y_i (number of events) for the i^{th} observation is assumed to obey the Poisson distribution, with λ_i representing both the mean and variance of the distribution for the i^{th} observation.

The relationship between the mean value of λ_i and the predictors is established by means of a log-linear model of the form

$$\log(\lambda_i) = \mathbf{x}_i^T \boldsymbol{\beta}, \quad (3)$$

where x_i is the vector-string of predictors for the i^{th} observation, β is the vector-column of regression coefficients.

This model allows estimating the influence of various factors on the failure rate of onboard equipment.

The basic Poisson distribution regression model is not always accurate enough to solve the research problems. To take into account non-linear dependencies between predictors and response, polynomial functions were added. These functions are created using polynomial transformers that expand the original feature vectors [13]. This approach improves the flexibility and accuracy of the model, especially when analyzing complex and nonlinear relationships. The use of polynomial features contributes to a better fit of the model to the original data and improves the accuracy of predictions, which is critical for predictive maintenance tasks.

The basic statements and assumptions when using polynomial features include:

1) polynomial transformation – for a given set of predictors $x_i = [x_{i1}, x_{i2}, ..., x_{ip}]$ polynomial features are formed by generating all possible polynomial combinations of predictors up to a given degree d;

2) extended feature vector – the initial feature vector x_i is augmented with polynomial terms, for example, for two predictors x_1 and x_2 at degree d = 2 the extended feature vector takes the following form

$$\mathbf{x}_i^{(\text{poly})} = \left[1, x_{i1}, x_{i2}, x_{i1}^2, x_{i1}x_{i2}, x_{i2}^2\right], \quad (4)$$

where the components of the vector include:

– constant (1);
– linear terms (x_{i1}, x_{i2});
– quadratic terms (x_{i1}^2, x_{i2}^2);
– product of predictors ($x_{i1}x_{i2}$).

The use of an extended feature vector allows to take into account non-linear types of interactions between predictors, which is especially important when analyzing complex systems such as aircraft on-board equipment.

The Poisson distribution regression model uses polynomial signs to account for complex relationships between predictors. The process begins by expanding the feature space, which involves transforming the original data to a higher dimension. The next step is to construct a logarithmic linear model taking into account the new features, and estimate the parameters using the maximum likelihood method (MLE). This approach allows taking into account more complex dependencies between variables, improves the accuracy of forecasts and has the following form

$$\log(\lambda_i) = \mathbf{x}_i^{(\text{poly})^T} \boldsymbol{\beta}_i, \quad (5)$$

where $\mathbf{x}_i^{(\text{poly})}$ is the extended polynomial feature vector.

To optimize computing resources and simplify the implementation, the mathematical model with polynomial features is limited to the second degree. Let us consider in more detail the mathematical description of the model with polynomial features of the second degree, which includes:

1. Initial vector of predictors

$$\mathbf{x}_i = \left[1, x_{i1}, x_{i2}\right]. \quad (6)$$

2. Extended polynomial feature vector (degree d = 2)

$$\mathbf{x}_i^{(\text{poly})} = \left[1, x_{i1}, x_{i2}, x_{i1}^2, x_{i1}x_{i2}, x_{i2}^2\right]. \quad (7)$$

3. Logarithmic linear model with polynomial features

$$\log(\lambda_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \\ + \beta_3 x_{i1}^2 + \beta_4 x_{i1} x_{i2} + \beta_5 x_{i2}^2. \quad (8)$$

4. Linear predictor expressed through polynomial signs

$$\eta_i = \log(\lambda_i) = \mathbf{x}_i^{(\text{poly})^T} \beta = \beta_0 + \beta_1 x_{i1} + \\ + \beta_2 x_{i2} + \beta_3 x_{i1}^2 + \beta_4 x_{i1} x_{i2} + \beta_5 x_{i2}^2. \quad (9)$$

5. Mean value (and at the same time dispersion) of the Poisson distribution

$$\lambda_i = \exp(\eta_i). \quad (10)$$

6. Logarithmic likelihood function for the Poisson distribution regression model logarithmic likelihood is defined by the following expression

$$\log L(\boldsymbol{\beta}_i) = \sum_{i=1}^{n} \left[y_i \log(\lambda_i) - \lambda_i - \log(y_i!)\right]. (11)$$

Substituting the expression for λ_i from equation (11) into equation (12), we obtain the expanded form of the logarithmic likelihood function of the form

$$\log\ L(\boldsymbol{\beta_i}) = \sum_{i=1}^{n}\left[y_i\left(\mathbf{x}_i^{(\text{poly})^T}\boldsymbol{\beta_i}\right) - \exp\left(\mathbf{x}_i^{(\text{poly})^T}\boldsymbol{\beta_i}\right) - \log\left(y_i!\right)\right]. \quad (12)$$

7. Maximum likelihood estimation (MLE), in which the coefficients β_i are estimated by maximizing the logarithmic likelihood function using a formula of the form

$$\hat{\boldsymbol{\beta}}_{\boldsymbol{i}} = \arg\max_{\beta}\log\ L(\boldsymbol{\beta_i}). \quad (13)$$

The maximum likelihood $\boldsymbol{\beta}$ method allows to find the optimal values of the coefficients β_i, which most fully correspond to the observed data, which provides a reliable basis for predicting the failure rate of onboard equipment, taking into account the nonlinear interactions between the factors.

Thus, the Poisson distribution regression model with polynomial signs allows to take into account non-linear dependencies between predictors and failure rate of onboard equipment. This approach is especially effective when analyzing complex technical systems, where interrelations between factors may be more complicated.

The choice of polynomial features of the second degree is conditioned by the desire to balance between increasing the accuracy of the model and refusal from retraining the program. Polynomials of the second degree allow to take into account nonlinear dependencies, while preserving the interpretability of the model.

Extending the feature space to second degree polynomials provides a compromise between the complexity of the model and its ability to reflect nonlinear interactions. This makes it possible to improve the accuracy of failure prediction without excessive complication of computational processes.

When applying this model to the tasks of forecasting the technical condition of aviation equipment, the following aspects should be taken into account:

1. Predictor selection – the parameters most relevant to assessing the reliability of specific systems and components must be carefully selected.

2. Interpretation of the coefficients – polynomial terms complicate the direct interpretation of the coefficients, so it is important to analyze their combined effect on the predicted failure rate.

3. Model validation – requires rigorous testing of the model against independent data to assess its predictive ability under real-world operating conditions.

4. Consideration of operational factors – the model should adapt to different equipment operating modes and aircraft operating conditions.

The application of the described methodology allows to create a flexible tool for predicting the reliability of onboard equipment, which in turn also contributes to the optimization of maintenance processes and improvement of flight safety.

The process of optimization of maintenance planning is based on the following criteria:

1. Minimization of total aircraft downtime.

2. Minimization of the intervals between maintenance forms.

3. Minimization of total maintenance costs.

4. Ensuring the required level of component reliability.

These criteria are taken into account in the objective function of the optimization procedure with appropriate weighting coefficients.

Data preparation

The data preparation process is a key step in the study to ensure the validity and correctness of the subsequent analysis. The methodology involves several steps of transforming raw statistical data on maintenance into a structured data set suitable for Poisson distribution regression analysis and predictive modeling.

1. Data Collection – complete maintenance logs and records of failure events for the annual period of operation of various aircraft components, from critical systems to support equipment are collected.

2. Data cleaning – a comprehensive processing of the original data set is carried out, including:

– identifying and eliminating outliers using statistical z-score analysis;

– processing missing values by interpolation methods based on nearby data points;

– correction of input errors and standardizing of component classification.

3. Data Transformation – continuous time-to-failure data is converted into discrete failure count intervals to prepare for Poisson regression.

4. Data Normalization – minimax normalization is carried out with respect to time-to-failure indicators to ensure comparability of data across different components, regardless of their usage.

5. Categorization – classification of components according to their criticality for aircraft operation, taking into account their functional role and the impact of failures on flight safety.

6. Generating the final dataset – the dataset prepared for the software is structured to reflect the failure cases of each component based on operating hours and operating conditions.

The result of this process is a comprehensive dataset that provides a robust basis for subsequent regression analysis of the Poisson distribution and predictive maintenance modeling. A sample dataset with different components is presented in Table 1.

Table 1 presents mean time between failures of the aircraft, number of failures, operating conditions and maintenance activities taken for the different component types during 2020, they provide a reliable basis for the presented analysis.

Preparation of data for modeling also includes:

1. Data cleaning to ensure the integrity and reliability of the dataset, including:

a) deviation detection and elimination, in which, using the inter-quartile range (IQR) method, deviations were identified and excluded from further analysis: mean time between failures and the number of failures that were 1.5 times greater the IQR from the quartiles;

Table 1

Dataset sample with diverse components

Component	MTBF	Number of failures	Operating conditions	Maintenance activities	Date of incident
Comp_A	3034.80	1	Extreme	Repair	2020-08-14
Comp_B	4369.73	3	Normal	Inspection	2020-04-08
Comp_B	710.00	3	Severe	Inspection	2020-07-19
Comp_B	2527.25	5	Normal	Repair	2020-12-25
Comp_C	907.73	2	Severe	Repair	2020-02-02
Comp_C	603.78	1	Extreme	Repair	2020-02-10
Comp_C	4924.54	3	Severe	Repair	2020-08-03
Comp_C	1547.47	4	Severe	Repair	2020-09-24
Comp_C	4881.90	3	Severe	Repair	2020-11-23
Comp_D	2221.08	0	Extreme	Repair	2020-03-02
Comp_D	2234.37	3	Normal	Inspection	2020-04-10
Comp_D	3561.38	1	Severe	Inspection	2020-06-20
Comp_D	559.69	5	Normal	Inspection	2020-08-01
Comp_D	3252.44	1	Severe	Replacement	2020-09-17
Comp_E	2861.49	3	Severe	Repair	2020-01-28
Comp_E	4739.91	5	Severe	Inspection	2020-02-04
Comp_E	3282.74	5	Extreme	Inspection	2020-02-17
Comp_E	2299.37	3	Severe	Repair	2020-05-14
Comp_E	531.80	4	Normal	Repair	2020-08-18
Comp_E	2600.43	1	Extreme	Inspection	2020-10-19

b) calculation of missing values, in which a linear interpolation method was used to provide a complete data set for analysis.

2. Data transformation for their adaptation and regression of the Poisson distribution, for which the following transformations were implemented:

a) The distribution of mean time between failures, with continuous mean time between failures distributed among cells to facilitate modeling of the number of failures within these intervals;

b) number of failures – was directly used as a response parameter in the Poisson distribution regression model, which met the requirements for its formation.

3. The data was normalized in the Python programming language using the Scipy library. The code is as follows

def min_max_normalize(data): return (data – data.min()) / (data.max() – data.min())

normalized_data = min_max_normalize(data).

This allowed the mean time between failures to be between 0 and 1, which allowed for a standardized scale for the different components.

4. Categorization of components by classification based on their criticality to aircraft operation, in particular:

a) critical components – engines and avionics that directly affect the safety and operational capabilities of the aircraft [14];

b) non-critical components – interior lighting and passenger entertainment systems whose failures have an insignificant impact on overall safety.

5. The final dataset was carefully assembled to ensure that all pre-processing steps were accurately reflected. This dataset formed the basis for a Poisson regression analysis and subsequent predictive maintenance modeling to provide insight into the failure patterns and maintenance needs of aircraft onboard equipment.

6. Application of Poisson distribution regression analysis to model aircraft component failure rates. The choice of Poisson distribution regression is due to its adequacy for analyzing count data typical for reliability studies.

The model is expressed as

$$\log(\lambda_i) = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_n X_{ni}, \qquad (14)$$

where λ_i is the expected failure rate of the component *i*, β_0, β_1, … β_n coefficients reflecting the influence of the variables X_1, …, X_n on the failure rate.

Stages of regression analysis:

1. Variable selection:

– the dependent variable is the number of failures;

– independent variables are operating hours, operating conditions, maintenance activities;

2. Data Coding:

– operating conditions – “Normal” = 0, “Severe” = 1, ‘Extreme” = 2;

– maintenance activities – “Inspection” = 0, “Repair” = 1, “Replacement” = 2.

3. The model is log (λ) = β_0 + β_1 × *Operating time* + β_2 × *Operating condition* + β_3 × *Maintenance measures;*

4. To estimate the model parameters, the maximum likelihood method is applied to estimate the coefficients β.

5. When interpreting the results, it is important to take into account that the analysis of β coefficients allows to evaluate the influence of each factor on the failure rate. A positive coefficient indicates an increase in the failure rate with the growth of the corresponding parameter.

The presented approach also provides a reliable basis for predicting the failure rate of onboard equipment and optimizing maintenance strategies.

Algorithm for optimization of maintenance planning

The main goal of optimization is to minimize equipment downtime and maintenance costs while ensuring the required level of reliability and flight safety [15].

Optimization Criteria:

1. The objective function is_qualitatively a combination of expected downtime due to failures and preventive maintenance costs:

– downtime cost (the product of the average downtime per failure event and the associated cost per hour of downtime);

– maintenance costs (the cost of each type of maintenance activities (inspection, repair, replacement)).

2. Restrictions:

– frequency of maintenance (established intervals (forms) of maintenance, taking into account technological schedules and availability of required resources);

– resource constraints (limits on the number of maintenance activities in a given period);

– regulatory requirements (compliance with safety and regulatory standards).

3. The genetic variant of the algorithm is chosen because it efficiently handles nonlinear problems containing many constraints.

The genetic algorithm generates and evaluates different variants of maintenance plans, iteratively improving their performance with respect to downtime and actual costs.

4. Analysis of the optimized maintenance process schedule involving evaluation of the resulting plan according to the criteria:

– efficiency (reduction of expected downtime and maintenance costs);

– feasibility (compliance with operational and regulatory requirements);

– improvement potential (identification of components or periods for further optimization).

In practice, based on a synthetic dataset, the optimization algorithm takes into account the failure rate of each component and maintenance statistics, for example, more frequent preventive maintenance may be recommended for components with high failure rates in harsh operating conditions.

The result of optimization is a maintenance plan that balances equipment reliability, operating costs and compliance with regulatory safety requirements.

The algorithm can show that product Comp_C that exhibits a high failure rate in harsh conditions significantly benefits from preventive replacement every 1000 hours of operation, reducing the total downtime by 20% compared to the existing schedule.

This detailed approach allows to determine how a genetic algorithm can be used to develop an optimal maintenance plan, significantly improving the efficiency of maintenance processes and equipment reliability based on comprehensive synthetic data analysis.

Model Adequacy Testing

To evaluate the effectiveness of the model, its predictions are compared with actual equipment failure data, using two main metrics:

– Root Mean Square Error (RMSE) – measures the accuracy of the predictions;

– Pearson Correlation Coefficient – measures the strength of the relationship between predicted and actual values. The following are provided:

Step 1. Cross-Validation Setup.

Cross-validation is used to evaluate the predictive performance of the model, where the dataset is divided into a training set (80%) used to develop the model and a test set (20%) used to evaluate its predictions. This division ensures that the model is tested on the data by simulating real-world forecasting scenarios.

Step 2. Model Performance Metrics.

Several metrics are used to quantify the accuracy and reliability of the model. These include:

1. Root Mean Square Error (RMSE), to measure the root mean square of prediction errors, providing an indication of the accuracy of the model by a formula of the following form

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}\left(y_i - \hat{y}_i\right)^2}, \qquad (15)$$

where y_i is the current failure value, $\widehat{y_i}$ is the predicted failure rate, and n is the number of observations on the test set.

2. Pearson correlation coefficient (r), which estimates the linear correlation between the actual and predicted failures rate using the formula of the following form

$$r = \frac{\sum_{i-1}^{n}\left(y_i - \bar{y}\right)\left(\hat{y}_i - \bar{\hat{y}}\right)}{\sqrt{\sum_{i-1}^{n}\left(y_i - \bar{y}\right)^2}\sqrt{\sum_{i-1}^{n}\left(\hat{y}_i - \bar{\hat{y}}\right)^2}}, \qquad (16)$$

where $\bar{y}$ and $\bar{\hat{y}}$ are the mean values of current and predicted failures, respectively.

Step 3. Applying the model to the test set.

The parameters of the Poisson distribution regression model are estimated on the training dataset, after which the model is applied to the test set to predict the number of failures. These predictions are then used to evaluate the performance of the model using the metrics defined above (Step 2).

Step 4. Evaluation of the optimization algorithm.

The effectiveness of the optimization algorithm is evaluated by implementing the optimized maintenance plan on a test plant and observing the resulting changes in failure rate and maintenance costs.

The optimized maintenance plan is compared to the actual plan used during the test period, assessing differences in performance and cost effectiveness.

Step 5. Statistical analysis to validate model predictions, including:

1. Significance testing using statistical tests such as the chi-squared test, which are used to determine whether differences between actual and predicted failure rates are statistically significant.

2. Calculating confidence intervals for Pearson correlation coefficients and root mean square error to quantify the uncertainty of model performance metrics.

Step 6. Validation.

The reliability of the results is assessed by analyzing the numerical performance metrics of the model. The main criteria include the root mean square error of prediction (RMSE), the Pearson correlation coefficient between actual and predicted values, as well as an assessment of the economic efficiency of the optimized maintenance plan.

Comparing model predictions with actual data allows to determine:

– accuracy of prediction failures of the aviation equipment product;

– effectiveness of the proposed maintenance strategy;

– the potential for further improvement of the model.

Regular updating of the model based on new operational data will ensure that its predictive ability is maintained with the required level of accuracy.

The results of the analysis are of key importance for improving the methodology of predicting the technical condition of aviation equipment products.

The application of the developed Poisson distribution regression models with polynomial features contributes to improving the accuracy of failure prediction, optimizing maintenance intervals and minimizing the risks of random failures of critical systems. This study makes a significant contribution to the development of the predictive maintenance concept in the aviation industry, which ultimately leads to improved flight safety and aircraft operational efficiency.

One of the key areas of applications of predictive models is the optimization of maintenance schedules. Accurate failure rate predictions help to determine the optimal timing of maintenance activities, ensuring that components are serviced before they fail. This proactive approach minimizes unexpected downtime and improves the efficiency of aircraft maintenance processes. By planning maintenance based on actual data rather than fixed intervals, airlines can reduce unnecessary maintenance costs and improve the overall reliability of their fleet.

The developed models facilitate decision making based on up-to-date operational data. This will allow maintenance planners to optimally allocate resources by focusing on critical components with a high probability of failure in the near future. This approach will improve maintenance efficiency and ensure timely maintenance of critical aircraft systems.

The article emphasizes the importance of model validating by comparing its predictions with actual maintenance statistics data. This verification process ensures the reliability and accuracy of the model predictions. Regularly updating the model using new operational data allows for continuous improvement of its predictive capabilities, leading to a constant improvement in the quality of maintenance planning and execution.

Thus, the results of the Poisson distribution regression and polynomial analysis provide useful information to help to plan maintenance, optimize resource allocation, reduce costs and improve the overall reliability and safety of aircraft operation. These models help to avoid the occurrence of unexpected events through a proactive maintenance strategy, ensuring that aviation products are maintained at the highest level of technical condition.

The dependencies presented in Figures 1–4 show the expected number of failures at a given operating time of products in the process of technical operation. The results are obtained by applying various methods of predicting the technical condition of products under expected operating conditions. Using each method, the mean square errors and Pearson correlation coefficients were estimated for the data set from Table 1.

Figure 1 shows a weak positive correlation between operating time and failure rate. The basic Poisson regression model shows a significant spread of predicted values relative to actual data, which is reflected in a high error (RMSE: 1.446) and low correlation (0.162).

Figure 2 shows the improvement in prediction accuracy due to the inclusion of nonlinear dependencies. The addition of polynomial features allowed the model to more accurately track changes in the number of failures at different operating hours, as evidenced by a decrease in RMSE to 1.146 and an increase in correlation to 0.240.

Figure 3 shows a complete coincidence of the predicted values with the actual data. The single-coded least squares method demonstrates unrealistically accurate prediction of the expected number of failures at any operating time (RMSE: 5.03e-14, correlation: 1.0), which indicates the need to change the data set for machine training of the model.

Figure 4 shows the results after removing the non-numerical parameters from the model. The dependence between operating time and number of failures becomes more realistic, with a moderate spread of predicted values (RMSE: 1.391, correlation: −0.319), which more fully corresponds to the real processes of technical operation.

Table 2 presents the results of the evaluation of the effectiveness of the proposed research methods. For approbation of the research results, the detailed mathematical formulation

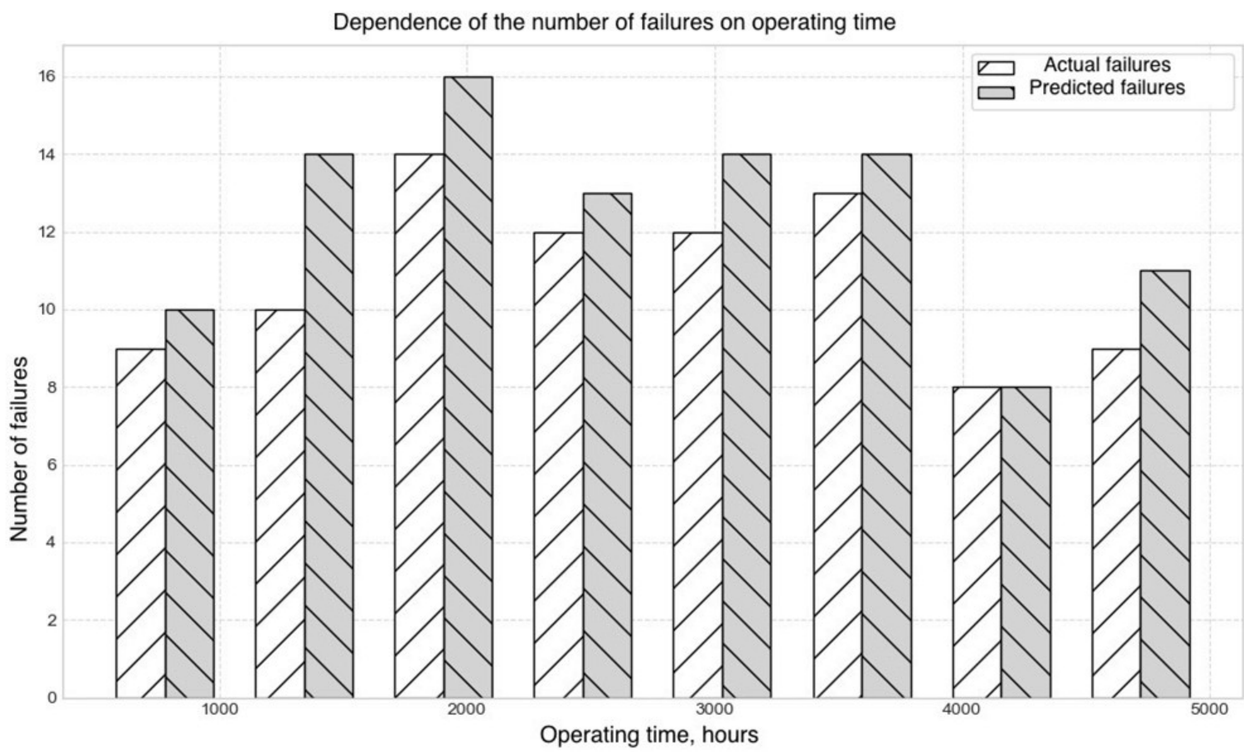

Fig. 1. Dependence of the number of failures on operating time using Poisson distribution regression. RMSE: 1.446000013304307; Pearson Correlation Coefficient: 0.1624522626478234

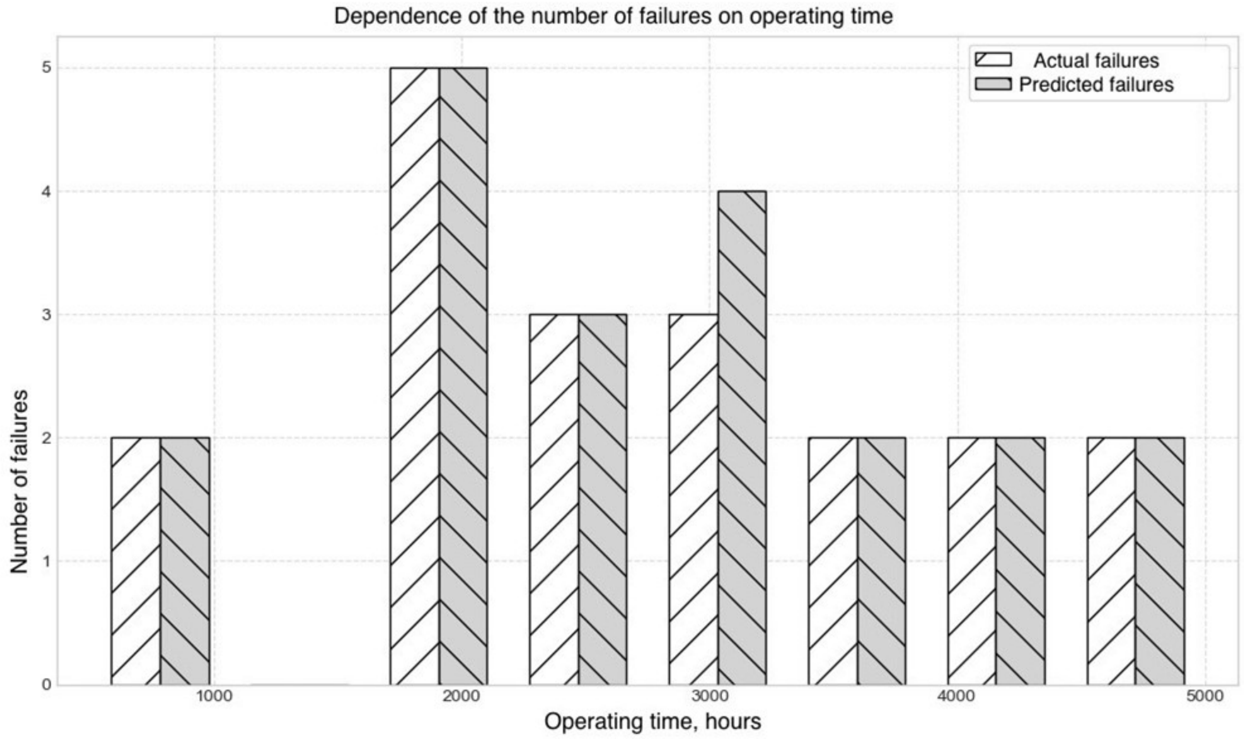

Fig. 2. Failures prediction using polynomial features in the Poisson distribution regression model.
RMSE: 1.1465954900923936; Pearson Correlation Coefficient: 0.240611370124739

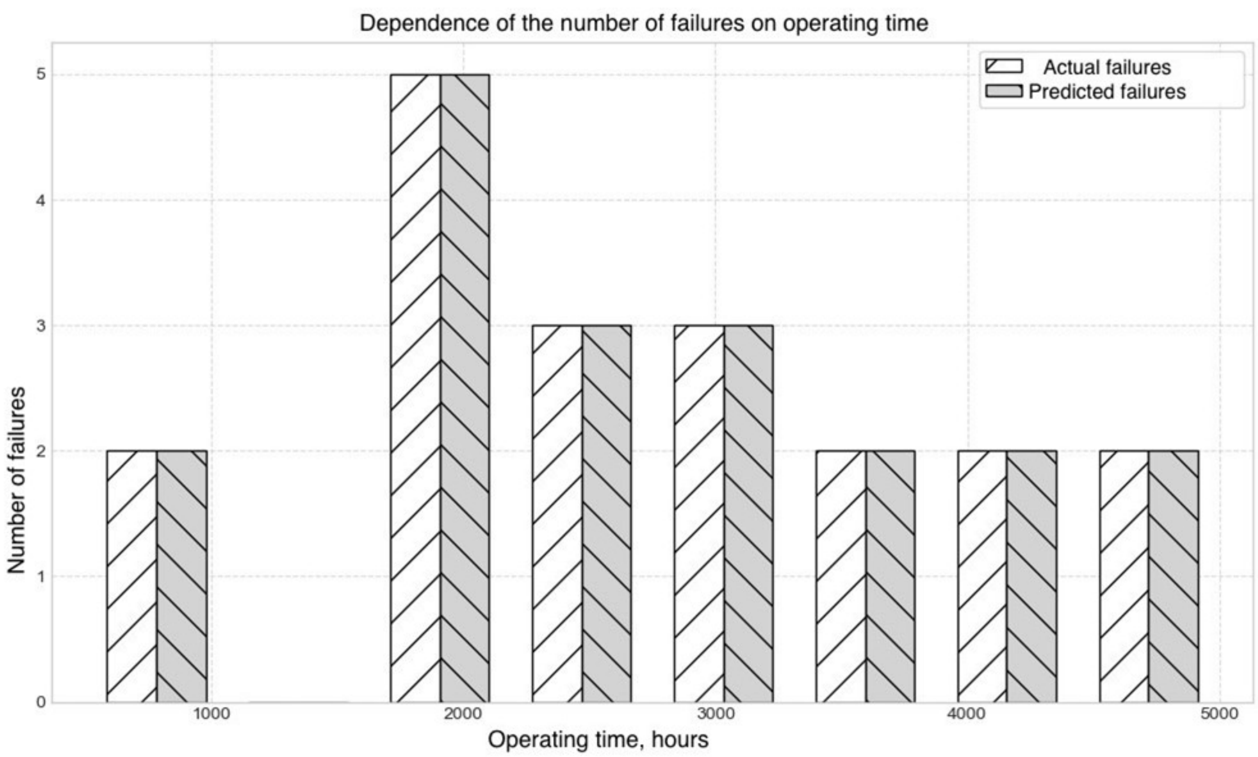

Fig. 3. Least squares simulation of failures with single encoding.
RMSE: 5.034734662940756e-14; Pearson Correlation Coefficient: 1.0

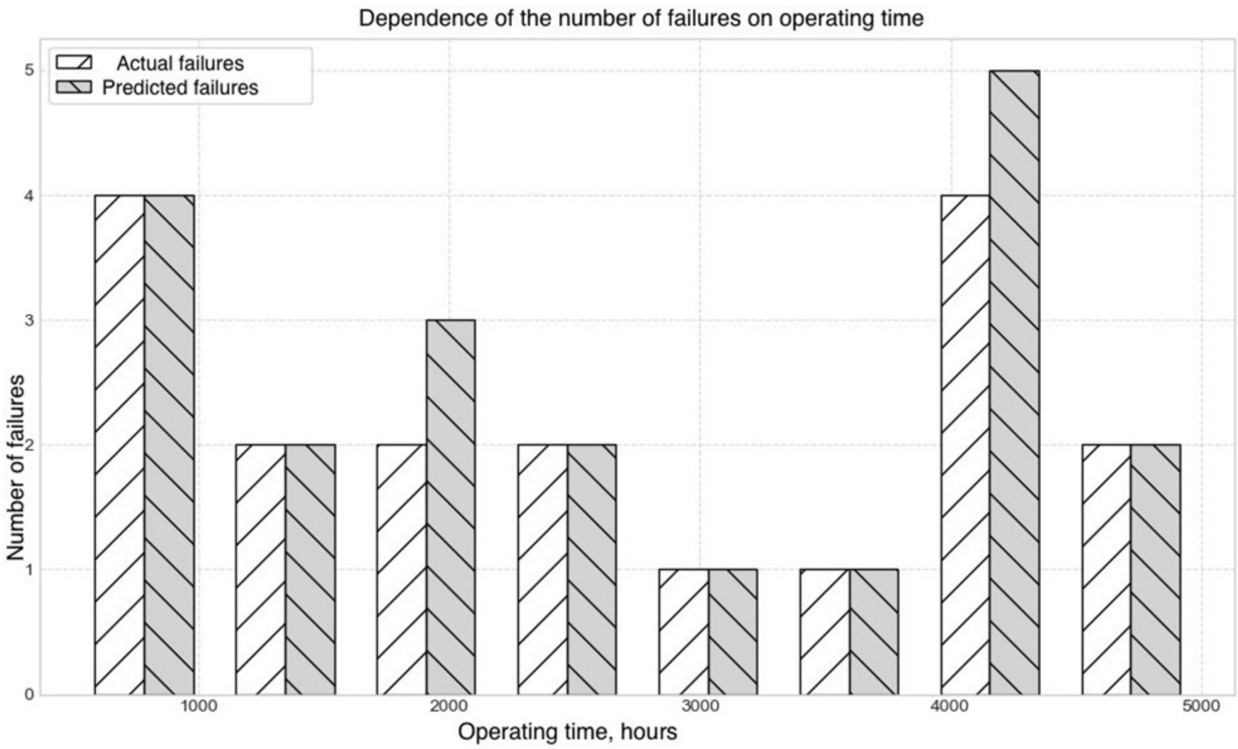

Fig. 4. The results of predicting failures using the least squares method after removing non-numeric parameters. RMSE: 1.3905425046960287; Pearson Correlation Coefficient: −0.3194427708423738

Table 2

Research methodology effectiveness

Code file	Methodology	Model used	RMSE	Pearson correlation coefficient
danyaplus2.py	Poisson distribution regression (fig. 1)	Generalized linear model	1.446000013304307	0.16245226264782342
danyaplus3.py	Polynomial signs (fig. 2)	Poisson distribution regression	1.1465954900923936	0.240611370124739
danyaplus4.py	Single coding (fig. 3)	Least squares regression	5.034734662940756e-14	1.0
danyaplus5.py	Single coding with removal of non-numeric columns (fig. 4)	Least squares regression	1.3905425046960287	−0.3194427708423738

of the proposed methods in the form of program codes has been placed in cloud storage.[1]

Comparative analysis of the results shows that the Poisson distribution regression method with polynomial features (danyaplus3.py, fig. 2) demonstrates the best balance between accuracy (RMSE = 1.146) and generalization ability (correlation coefficient = 0.240). The single-coded least squares method (danyaplus4.py, fig. 3)

[1] Detailed mathematical formulation of the proposed methods in the form of Python program codes. *Yandex disk.* Available at: https://disk.yandex.ru/d/1lUjYl2SX4nmug (accessed: 02.04.2025). (in Russian)

shows suspiciously perfect results, which may indicate model overfitting.

Abnormally high results for the least squares method with single coding (RMSE ≈ 5.03e-14, correlation coefficient = 1.0) indicate probable overfitting of the model. This may be due to the fact that the model has adjusted too accurately to the features of the synthetic data, losing its generalization ability. In real conditions, such results are unlikely and require additional verification on an array of independent data.

Conclusion

1. Quantitative analysis of the performance of the prediction methodologies showed significant differences in the accuracy of the models. The basic regression of Poisson distribution demonstrated RMSE = 1.446 and Pearson correlation coefficient equal to 0.162. The introduction of polynomial signs led to improvement of indicators: RMSE decreased to 1.146, correlation coefficient increased to 0.240, which confirms the effectiveness of polynomial signs application for modeling of nonlinear dependencies in the processes of technical operation of an aviation product.

2. The single-coded least-squares regression method showed statistically abnormal results (RMSE ≈ 5.03e-14, correlation coefficient = 1.0), indicating the overfitting of the model on synthetic data. This effect requires the implementation of regularization and cross-validation methods to improve the generalizing ability of the model.

3. Application of the Poisson distribution regression model with polynomial signs of the second degree provides an optimal balance between the complexity of the model and its ability to reflect nonlinear interactions in the processes of technical operation of the aviation product, which is confirmed by the improvement of prediction accuracy indicators.

4. The main limitation of this study is the use of synthetic data, which does not fully reflect the complexity and variability of real aircraft maintenance processes. In particular, synthetic data do not take into account all possible anomalies and rare cases of failures, which may lead to distortion of modeling results.

5. The results of the study confirm the effectiveness of using Poisson distribution regression with polynomial signs for predicting failures of aviation product. At the same time, the revealed limitations of synthetic data indicate the necessity of model validation on real operational data.

6. The practical significance of the developed models lies in the possibility of their integration into the aircraft maintenance program to predict failures of aviation product and optimize maintenance and repair works, which potentially contributes to the reduction of operating costs and improvement of flight safety.

7. To overcome the identified limitations, it is further necessary to:

– Conduct model validation on real operational data;

– Implement regularization methods to prevent overfitting;

– Develop mechanisms for model adaptation to different aviation product types and operating conditions;

– Explore the possibilities of integrating additional factors into the model to improve prediction accuracy.

References

1. **Aleksanyan, A.R., Itskovich, A.A., Faynburg, I.A.** (2014). The method of integrated logistics support when forming procedures for maintenance of aircraft airworthiness. *Nauchnyy vestnik MGTU GA*, no. 205, pp. 22–27. (in Russian)

2. **Sizikov, D.O.** (2024). Method of time series analysis and its mathematical model in software. *International Research Journal*, no. 3 (141). DOI: 10.23670/IRJ.2024.141.5 (accessed: 02.04.2025).

3. **Itskovich, A.A., Faynburg, I.A., Aleksanyan, A.R.** (2015). Methodological aspects of software processes management of the airworthiness of aircraft. *Nauchnyy vestnik MGTU GA*, no. 219, pp. 12–19. (in Russian)

4. **Kuznetsov, S.V.** (2015). Processes and systems of avionics technical operation as Markov and Semimarkov processes mathematical

models. *Nauchnyy vestnik MGTU GA*, no. 213, pp. 28–33. (in Russian)

5. Kuznetsov, S.V. (2017). Avionics technical operation system and scientific basis for its formation. *Civil Aviation High Technologies*, vol. 20, no. 6, pp. 15–24. DOI: 10.26467/2079-0619-2017-20-6-15-24 (in Russian)

6. Bukirev, A.S., Savchenko, A.Y., Yatsechko, M.I., Malyshev, V.A. (2020). Diagnostic system for the technical condition of the aircraft avionics complex based on intelligent information technologies. *Modeling, Optimization and Information Technology*, vol. 8, no. 1 (28). DOI: 10.26102/2310-6018/2020.28.1.010 (accessed: 02.04.2025). (in Russian)

7. Bühlmann, P., Van De Geer, S. (2011). Statistics for high-dimensional data: methods, theory and applications. Springer Science & Business Media, 558 p.

8. Zasukhin, A.S., Budaev, V.D., Sizikov, D.O., Sadegzade, S.H. (2021). Evaluating two navigation systems error deviation algorithm. *Vestnik of Russian new university. Series: complex systems: models, analysis, management,* no. 4/1, pp. 3–13. DOI: 10.18137/RNU.V9187.21.04/1.P. 003 (in Russian)

9. Kuznetsov, S.V. (2023). Characteristics of operational control reliability of aircraft functional systems and avionics suites. *Civil Aviation High Technologies*, vol. 26, no. 6, pp. 58–74. DOI: 10.26467/2079-0619-2023-26-6-58-74 (in Russian)

10. Ying, X. (2019). An overview of overfitting and its solutions. *Journal of physics: Conference series*, vol. 1168, issue 2, ID: 022022. DOI: 10.1088/1742-6596/1168/2/022022 (accessed: 02.04.2025).

11. Montesinos López, O.A., Montesinos López, A., Crossa, J. (2022). Overfitting, model tuning, and evaluation of prediction performance. In book: Multivariate Statistical Machine Learning Methods for Genomic. Cham: Springer International Publishing, pp. 109–139. DOI: 10.1007/978-3-030-89010-0_4

12. Muñoz-Pichardo, J.M., Pino-Mejías, R., García-Heras, J., Ruiz-Muñoz, F., González-Regalado, M.L. (2021). A multivariate Poisson regression model for count data. *Journal of Applied Statistics*, vol. 48, no. 13–15, pp. 2525–2541. DOI: 10.1080/02664763.2021.1877637

13. Heaton, J. (2016). An empirical analysis of feature engineering for predictive modeling. *In: SoutheastCon 2016*. Norfolk, VA, USA, 6 p. DOI: 10.1109/SECON.2016.7506650 (accessed: 02.04.2025).

14. Fourar, H.E., Ogunvoul, B.D., Budaev, V.D., Lachi, F. (2020). Risk assessment in ensuring aircraft on-board systems safety. *Civil Aviation High Technologies*, vol. 23, no. 4, pp. 84–95. DOI: 10.26467/2079-0619-2020-23-4-84-95 (in Russian)

15. Budaev, V.D. (2024). Methods of optimization of aircraft fleet maintenance planning. *International Research Journal*, no. 7 (145). DOI: 10.60797/IRJ.2024.145.6 (accessed: 02.04.2025).

Список литературы

1. Алексанян А.Р., Ицкович А.А., Файнбург И.А. Интегрированная логистическая поддержка формирования процедур поддержания летной годности воздушных судов // Научный вестник МГТУ ГА. 2014. № 205. С. 22–27.

2. Сизиков Д.О. Метод анализа временных рядов и его математическая модель в программном обеспечении [Электронный ресурс] // Международный научно-исследовательский журнал. 2024. № 3 (141). DOI: 10.23670/IRJ.2024.141.5 (дата обращения: 02.04.2025).

3. Ицкович А.А., Файнбург И.А., Алексанян А.Р. Методологические аспекты программного управления процессами поддержания летной годности воздушных судов // Научный вестник МГТУ ГА. 2015. № 219. С. 12–19.

4. Кузнецов С.В. Математические модели процессов и систем технической эксплуатации авионики как марковские и полумарковские процессы // Научный вестник МГТУ ГА. 2015. № 213 (3). С. 28–33.

5. Кузнецов С.В. Система технической эксплуатации авионики и научные основы ее формирования // Научный вестник МГТУ ГА. 2017. Т. 20, № 6. С. 15–24. DOI: 10.26467/2079-0619-2017-20-6-15-24

6. Букирев А.С. Система диагностики технического состояния комплекса бортового оборудования воздушного судна на основе интеллектуальных информационных технологий / А.С. Букирев, А.Ю. Савченко, М.И. Яцечко, В.А. Малышев [Электронный ресурс] // Моделирование, оптимизация и информационные технологии. 2020. Т. 8, № 1 (28). DOI: 10.26102/2310-6018/2020.28.1.010 (дата обращения: 02.04.2025).

7. Bühlmann P., Van De Geer S. Statistics for high-dimensional data: methods, theory and applications. Springer Science & Business Media, 2011. 558 p.

8. Засухин А.С. Исследование алгоритма оценки погрешности отклонения двух навигационных систем / А.С. Засухин, В.Д. Будаев, Д.О. Сизиков, С.Х. Садегзаде // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. 2021. № 4/1. С. 3–13. DOI: 10.18137/RNU.V9187.21.04/1.P. 003

9. Кузнецов С.В. Характеристики достоверности эксплуатационного контроля функциональных систем и комплексов бортового оборудования воздушных судов // Научный вестник МГТУ ГА. 2023. Т. 26, № 6. С. 58–74. DOI: 10.26467/2079-0619-2023-26-6-58-74

10. Ying X. An overview of overfitting and its solutions [Электронный ресурс] // Journal of physics: Conference series. 2019. Vol. 1168, iss. 2. ID: 022022. DOI: 10.1088/1742-6596/1168/2/022022 (дата обращения: 02.04.2025).

11. Montesinos López O.A., Montesinos López A., Crossa J. Overfitting, model tuning, and evaluation of prediction performance // Multivariate Statistical Machine Learning Methods for Genomic. Cham: Springer International Publishing, 2022. Pp. 109–139. DOI: 10.1007/978-3-030-89010-0_4

12. Muñoz-Pichardo J.M. A multivariate Poisson regression model for count data / J.M. Muñoz-Pichardo, R. Pino-Mejías, J. García-Heras, F. Ruiz-Muñoz , M.L. González-Regalado // Journal of Applied Statistics. 2021. Vol. 48, no. 13–15. Pp. 2525–2541. DOI: 10.1080/02664763.2021.1877637

13. Heaton J. An empirical analysis of feature engineering for predictive modeling [Электронный ресурс] // SoutheastCon 2016. Norfolk, VA, USA, 2016. 6 p. DOI: 10.1109/SECON.2016.7506650 (дата обращения: 02.04.2025).

14. Фурар Х.Э. Оценка рисков при обеспечении безопасности бортовых систем воздушного судна / Х.Э. Фурар, Б.Д. Огунвоул, В.Д. Будаев, Ф. Лаши // Научный вестник МГТУ ГА. 2020. Т. 23, № 4. С. 84–95. DOI: 10.26467/2079-0619-2020-23-4-84-95

15. Будаев В.Д. Методы оптимизации планирования технического обслуживания авиационного парка [Электронный ресурс] // Международный научно-исследовательский журнал. 2024. № 7 (145). DOI: 10.60797/IRJ.2024.145.6 (дата обращения: 02.04.2025).

Information about the authors

Blessing I. Ogunvoul, Candidate of Technical Sciences, Associate Professor, Flight and Life Safety Chair, Moscow State Technical University of Civil Aviation, ogunvouluni@yandex.ru.

Vladislav D. Budaev, Senior Lecturer, the Chair of Aircraft Engine Engineering, Moscow State Technical University of Civil Aviation, vlad_budaev@mail.ru.

Daniil O. Sizikov, Senior Lecturer, Electrical Systems and Flight Navigation Complexes Maintenance Chair, Moscow State Technical University of Civil Aviation, d.sizikov@mstuca.ru.

Nikita V. Gorbakon, Senior Lecturer, the Chair of Aircraft Engine Engineering, Moscow State Technical University of Civil Aviation, n.gorbakon@mstuca.ru.

Arusya V. Vlasova, Candidate of Technical Sciences, Associate Professor, Transportation Organization on Air Transport Chair, Moscow State Technical University of Civil Aviation, a.vlasova@mstuca.ru.

Сведения об авторах

Огунвоул Блессинг Израилевич, кандидат технических наук, доцент кафедры безопасности полетов и жизнедеятельности МГТУ ГА, ogunvouluni@yandex.ru.

Будаев Владислав Дмитриевич, старший преподаватель кафедры двигателей летательных аппаратов МГТУ ГА, Vlad_budaev@mail.ru.

Сизиков Даниил Олегович, старший преподаватель кафедры технической эксплуатации авиационных электросистем и пилотажно-навигационных комплексов МГТУ ГА, d.sizikov@mstuca.ru.

Горбаконь Никита Вадимович, старший преподаватель кафедры двигателей летательных аппаратов МГТУ ГА, n.gorbakon@mstuca.ru.

Власова Аруся Витальевна, кандидат технических наук, доцент кафедры организации перевозок на воздушном транспорте МГТУ ГА, a.vlasova@mstuca.ru

Поступила в редакцию	30.06.2024	Received	30.06.2024
Одобрена после рецензирования	05.01.2025	Approved after reviewing	05.01.2025
Принята в печать	25.03.2025	Accepted for publication	25.03.2025