BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 73(6), 2025, Article number: e155046
DOI: 10.24425/bpasts.2025.155046

MECHANICAL AND AERONAUTICAL ENGINEERING, THERMODYNAMICS

Research on the causes of erroneous forecast generated using the Chapman-Kolmogorov equations for the process of vehicle operation

Jarosław ZIÓŁKOWSKI*, **Aleksandra LĘGAS**, **Mateusz OSZCZYPAŁA**, **Jakub KONWERSKI**, and **Jerzy MAŁACHOWSKI**

Faculty of Mechanical Engineering, Military University of Technology, ul. Gen. Sylwestra Kaliskiego 2, 00-908 Warsaw, Poland

Abstract. The work attempts to investigate the causes of incorrect predictions of the Chapman-Kolmogorov system of equations generated during vehicle operation. When researching the process of exploitation of technical objects, Markov theory is often used in literature on the subject. Based on the developed Markov or semi-Markov models, on the one hand, basic reliability indicators (such as readiness) are assessed, and the evolution of the considered operation process is anticipated. The solutions of the Chapman-Kolmogorov system serve as the basis for preparing the forecast. For applications, forecasts of limit probabilities, determination times, and oscillation parameters of the probabilities of the states of the exploitation process are useful. The literature on the subject indicates the interdependence of each forecast on the estimation errors of all elements of the transition intensity matrix of the model, as well as errors in the calculation of its eigenvalues, as a potential cause of unsatisfactory forecast performance in continuous time. Considering the above, the main topic of this work was to investigate the correctness of the Chapman-Kolmogorov assumption for the vehicle operation process, the solution of which will make a significant substantive contribution to the current state of knowledge on modeling operation processes.

Keywords: Markov model; forecast; Chapman-Kolmogorov equations; vehicle operation.

1. INTRODUCTION

The issues related to modelling the operation of various objects [1–3] are often discussed in the source literature. The authors mainly focus on reproducing the current process and the characteristic features of the tested object, or trials of its improvement [4]. The aspects mentioned above are generally necessary for proper control [5] of the process of operation, considering the objectives and tasks executed by the tested object. Scientific studies in this field can be divided into two main groups:

- The first one examines the current status (the present) of the phenomenon under investigation.
- The second one focuses on forecasting the evolution of the operation process (the oriented future).

The first group includes studies examining the current state of the process under investigation, in which authors analyze the basic reliability indicators (including readiness) and highlight economic and/or ecological aspects to make rational decisions and optimally benefit from the utility values of the object. For this purpose, the studied process is modeled using neural networks, reliability and renewal theory, as well as experimental, simulation, and operational studies [6–9] (including optimization) and event-based models that preserve trend and seasonality. Stochastic processes are commonly used to describe the phenomenon under study [10], including Markov theory. In the source literature, neural networks are used as practical tools for optimizing energy consumption [11], performing Pareto estimations [12], or detecting errors in batch production [13]. Reliability and renewal theory are widely used to describe operations [14, 15], considering their basic processes, i.e., use [16, 17] and maintenance [18–20]. Analyses of reliability [21–25], readiness [26–28], and availability are of high importance in this context [29, 30]. Some studies focus on the determination of key operational parameters [31–33] or refer to costs [34, 35].

The use of analytical methods to analyze and evaluate complex operational processes requires the development of detailed databases [36, 37]. The variability of the operating environment, as well as the dynamics of the process itself, complicate or even hinder the resolution of complex decision-making problems. In such situations, simulation methods and techniques are often used, where the basis of the simulation model is formed by empirical probability distributions of random variables. For example, in their paper [38], Izdebski *et al.* present the problem of minimizing the risk of hazardous events on the apron with the participation of aircraft and ground handling vehicles. An evaluation of the efficiency of airport processes using simulation tools is presented in the paper [39]. A methodology for the analysis and evaluation of the selected indices of the helicopter readiness used in the Polish Navy is considered in [40]. In turn, operational research is used to optimize the analyzed process, for example, to determine the optimal route [41], the minimum delivery time

*e-mail: jaroslaw.ziolkowski@wat.edu.pl

Manuscript submitted 2025-02-27, revised 2025-07-23, initially accepted for publication 2025-08-06, published in November 2025.

in the product distribution network [42], or to conduct customer satisfaction surveys [43].

To forecast the evolution of technical object operation processes over time using stochastic process models [44–46], Markov and semi-Markov models are often used [47–51], and the evolution of the operation process under consideration is predicted based on the solutions of the Chapman-Kolmogorov equations [52]. In terms of application, forecasts of limit probabilities, determination times, and oscillation parameters of the probabilities of the operation process states are of particular importance. A satisfactory level of forecast error was achieved for discrete time [53], while for continuous time, the level of forecast error for characteristic times was unsatisfactory, and unreasonable limit probabilities of process states, such as those for urban agglomerations, were found, calculated according to the Chapman-Kolmogorov system of equations. As a potential cause of the unsatisfactory level of forecast error [54, 55] in continuous time, the interdependence of each forecast on the estimation errors of all elements of the transition intensity matrix of the model and the calculation errors of its eigenvalues is indicated. This results from the definition of eigenvalues, which are solutions of a matrix equation containing all elements of the transition intensity matrix and their linear combinations and/or power functions that determine the solutions of the Chapman-Kolmogorov system of equations.

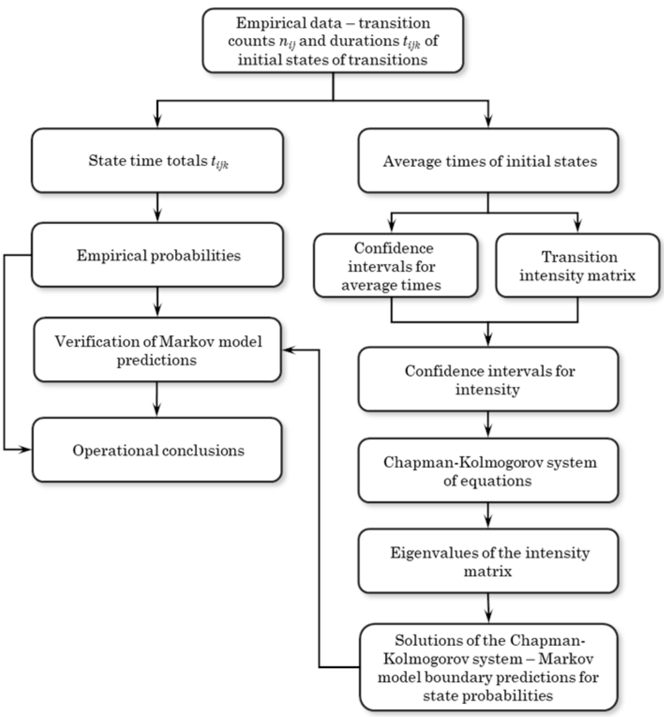

Fig. 1. Factors determining system forecasts Chapman-Kolmogorov according to [56]

For Markov models with more than four analytical states, investigations into forecast errors of the Chapman-Kolmogorov system of equations are executable only for special forms of the transition intensity matrix, when analytical formulas for their eigenvalues exist, or when the analytical solutions of the Chapman-Kolmogorov system of equations are not overly complex.

For the general form of the transition intensity matrix of degrees higher than four, only numerical studies of the forecast errors of the Chapman-Kolmogorov system of equations are feasible. These studies are conducted at selected points in the set of confidence intervals for elements of the transition intensity matrix, for example, at the centers and endpoints. Due to the considerable number of confidence intervals, numerical tests may prove to be either unreliable with insufficient test points or even infeasible within an acceptable time. For example, for three points of each confidence interval and M intervals, there are as many Chapman-Kolmogorov systems of equations that need to be solved as the number of three-element variations with repetitions from the set M of elements (M^3). For the process of operation considered in this work, with the number of intervals $M = 16$, it would be necessary to solve 4096 Chapman-Kolmogorov systems of equations for five probabilities of states. This would require the development of a special computational procedure and several hours of operation on an efficient computer, with low reliability of the calculated maximum forecast errors of the Chapman-Kolmogorov system of equations. A more reliable study would require calculations of forecasts at a minimum of ten points of each confidence interval. Therefore, for the process of operation considered below, $16^{10} \approx 1.1 \cdot 10^{12}$ Chapman-Kolmogorov systems of equations would need to be solved. The time needed to solve 10 000 systems using the Wolfram computer available "in the cloud" for Mathematica owners is 1.8 seconds. Calculating solutions for $1.1 \cdot 10^{12}$ systems would take over six years, and its cost would exceed PLN 220 million. Based on the above examples, it can be concluded that unacceptably high time and cost make it currently impossible to perform reliable numerical tests of forecast errors of the Chapman-Kolmogorov system of equations, even for only a five-state operation process, and the calculations using efficient computers would require verification of reliability. Such verification would require the use of gradient numerical methods.

The set of potential causes of forecast errors in the Chapman-Kolmogorov system of equations includes the following:

- Multi-seasonality of everyday life as the environment of operational processes [57]
- Autocorrelations, correlations, and heteroskedasticity of stochastic model variables [58]
- Deviations of the actual distributions of process state durations from the theoretical distributions

The fundamental reason may also be the failure to meet the Chapman-Kolmogorov assumption about the linear dependence of the derivatives of probabilities of states with respect to time on the values of their probabilities. The applicability of this assumption to operational process models has not been tested so far [59]. For this reason, examining the correctness of the Chapman-Kolmogorov assumption for the vehicle operation process was considered the main topic of the present work, making an important contribution to the current state of knowledge about modeling operation processes.

The paper is organized as follows. Section 1 contains a summary of knowledge about modeling the operation process, along with the justification for the purposefulness of the present work. Section 2 presents a multi-stage methodology for testing the Chapman-Kolmogorov system of equations, including initial conditions and the normalization condition. The results of the tests conducted on the actual operation process are described in Section 3, while Section 4 contains the results of the validation of the verification procedure for the Chapman-Kolmogorov assumptions. The verification procedure included: average frequencies and *ex ante* forecasts for average state probabilities, endpoints of confidence intervals for transition intensities, examination of nonlinearity of multiple and simple regression models, and validation of the proposed verification procedure based on simulations.

2. RESEARCH METHODOLOGY

Let $(\Omega,\ F,\ P)$ be a probability space. The sequence $\{X_t\}_{t\geq 0}$ denotes a continuous-time Markov chain, where $X_t\colon\ \Omega \longrightarrow \{S_1, S_2, S_3, S_4, S_5\}$. The distribution of the sojourn time, among other characteristics, is defined accordingly.

For a five-state stochastic process, the Chapman-Kolmogorov system of equations with the normalization condition and initial conditions has the following matrix form:

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}p_i\right] = \left[\lambda_{ij}\right]\cdot\left[p_i\right], \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_1 = \lambda_{11}p_1+\lambda_{12}p_2+\lambda_{13}p_3+\lambda_{14}p_4+\lambda_{15}p_5\,, \tag{1.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_2 = \lambda_{21}p_1+\lambda_{22}p_2+\lambda_{23}p_3+\lambda_{24}p_4+\lambda_{25}p_5\,, \tag{1.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_3 = \lambda_{31}p_1+\lambda_{32}p_2+\lambda_{33}p_3+\lambda_{34}p_4+\lambda_{35}p_5\,, \tag{1.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_4 = \lambda_{41}p_1+\lambda_{42}p_2+\lambda_{43}p_3+\lambda_{44}p_4+\lambda_{45}p_5\,, \tag{1.4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_5 = \lambda_{51}p_1+\lambda_{52}p_2+\lambda_{53}p_3+\lambda_{54}p_4+\lambda_{55}p_5\,. \tag{1.5}$$

Normalization condition

$$p_1+p_2+p_3+p_4+p_5 = 1. \tag{1.6}$$

The initial conditions

$$\exists_k\,(p_k=1)\cap\forall_{i\neq k}\,(p_{i\neq k}=0), \tag{1.7}$$

where $t\geq 0$ – non-negative continuous physical time of the process from the initial state S_3; $[p_i]$ – columnar probability vector p_1 of the process states; $\left[\frac{\mathrm{d}}{\mathrm{d}t}p_i\right]$ – column vector of time derivative probabilities t; $p_1 = p_1(t)$, $p_2 = p_2(t)$, $p_3 = p_3(t)$, $p_4 = p_4(t)$, $p_5 = p_5(t)$ – unknown probability of the process being in the states S_1–S_5 as functions of time t; $[\lambda_{ij}]$ – square transition intensity matrix, with elements representing transitions from state S_i to state S_j (rows and columns indexed by i and j); $\lambda_{ij\neq i}$ – off-diagonal elements of the intensity matrix, representing the reciprocals of the expected durations in state S_i before transitioning to S_j, or zeros if the transition probability from S_i to S_j is zero; λ_{ii} – diagonal elements of the intensity matrix, which are non-positive (negative or zero), defined such that the sum of each row (i.e., the diagonal and off-diagonal elements in row i) equals zero.

The Chapman-Kolmogorov assumption refers to equations (1.1)–(1.5), which constitute a homogeneous system of first-order ordinary differential equations, having infinitely many general solutions. Particular solutions are obtained after supplementing the system with the normalization condition for state probabilities (1.6) and the set of initial conditions (1.7), with one $p_i(t=0^+) = 1$.

The Chapman-Kolmogorov assumption is formally equivalent to the assumption of the multiple linear regression model for derivatives of probabilities with respect to time t, without free (constant) terms. This is also equivalent to the assumption of no heteroscedasticity in the model with respect to time. There is a well-known methodology and available software in literature that can perform estimation and post-estimation analyses of regression models. The slopes of the independent variables can be estimated from the sample, along with their errors and statistical significance, using multiple methods. This enables an empirical, direct verification of the Chapman-Kolmogorov assumption without solving the Chapman-Kolmogorov system of equations, by comparing the transition intensity estimates from the regression equations with the values calculated from the sample according to (4). A detailed verification of the Chapman-Kolmogorov assumption, with identification of the probabilities that do not meet the assumptions, is possible using the four-step procedure discussed below. It constitutes an indirect proof of the thesis known from logic, consisting in demonstrating the contradiction between the Chapman-Kolmogorov assumption and empirical data.

Stage 1 concerns the formulation of the empirical analog of differential equations (1.1)–(1.5) in the conventional form of an OLS-estimated multiple linear regression model without a constant term, to fulfil normalization condition (1.6) as $t\longrightarrow\infty$. The model includes estimators of p_i, $\frac{\mathrm{d}}{\mathrm{d}t}p_i$, and λ_{ij} as in (1), and has the following general form shown in (2), with detailed forms given in (2.1)–(2.5).

$$r_i(t) = \frac{\Delta w_i(t)}{\Delta t} = \frac{w_i(t)-w_i(t_0)}{t-t_0} = \sum_j a_{ij}w_j(t), \tag{2}$$

$$r_1 = a_{11}w_1+a_{12}w_2+a_{13}w_3+a_{14}w_4+a_{15}w_5\,, \tag{2.1}$$

$$r_2 = a_{21}w_1+a_{22}w_2+a_{23}w_3+a_{24}w_4+a_{25}w_5\,, \tag{2.2}$$

$$r_3 = a_{31}w_1+a_{32}w_2+a_{33}w_3+a_{34}w_4+a_{35}w_5\,, \tag{2.3}$$

$$r_4 = a_{41}w_1+a_{42}w_2+a_{43}w_3+a_{44}w_4+a_{45}w_5\,, \tag{2.4}$$

$$r_5 = a_{51}w_1+a_{52}w_2+a_{53}w_3+a_{54}w_4+a_{55}w_5\,. \tag{2.5}$$

where i – number of the empirical Chapman-Kolmogorov equation and the dependent variable r_i; j – number of the regressor (independent variable) $w_j(t)$; $t\geq 0$ – continuous time of the process; $\Delta t = t - t_0$ – time increment between times $t_0 < t$ and t; $t_0 < t$ – reference time of the increment Δt; min $\Delta t = 1$ minute (the accuracy of process clock), or the time interval of averaging large frequency fluctuations illustrated on Fig. 4 below, up to 40 000 minutes (15–20 process steps); $\Delta w_i(t) = w_i(t) - w_i(t_0)$ – the increment of frequency w_i in time interval $\langle t_0; t_0 + \Delta t \rangle$; $w_i(t)$, $w_j(t)$ – the cumulative frequencies in time period $\langle 0; t \rangle$ (estimators of $p_i(t)$, $p_j(t)$ in (1)); $r_i(t)$ – instantaneous rate of frequency change in time t is the estimator of derivative $\frac{\mathrm{d}}{\mathrm{d}t} p_i$; a_{ij} – directional coefficient (slope) of the model (2) for i-th equation and j-th regressor.

According to auxiliary system (2), the value a_{ij} denotes λ_{ij} intensity estimator in the Chapman-Kolmogorov system (1). The frequency $w_i(t)$ in continuous time is defined by (11). The conventional one-time estimator λ^*_{ij} is the reciprocal of the mean duration of the state before transition.

For an empirical phase trajectory, conditions (1.6) and (1.7) are satisfied by formula (11).

Structural and stochastic parameters, as well as confidence intervals for structural parameters, are calculated (estimated) based on empirical difference quotient values of $r_j(t)$ and frequencies $w_i(t)$, $w_j(t)$. Selecting the minimum time difference $\Delta t = t - t_0$ should consider the time needed for the examined process to establish itself satisfactorily in discrete time, to avoid the lack of observations of most process states in the time interval $\langle t - t_0; t \rangle$, and frequent jumps of $r_i(t)$ from 0.

Stage 2 of the procedure for verifying the Chapman-Kolmogorov assumption involves the estimation of structural and stochastic parameters, as well as the confidence intervals of structural parameters of the linear regression models defined by equation (2) for all states of the analyzed process. The structural parameters of models (2.1)–(2.5) are estimators of the coefficients in equations (1.1)–(1.5). The estimation is performed using various analytical programs, such as the free software Gretl.

Stage 3 of the procedure for verifying the Chapman-Kolmogorov assumption is the estimation of the elements of the transition intensity matrix from the data sample, and calculation of estimation errors and the endpoints of confidence intervals for transition intensities, according to the following formulas:

$$\lambda^{\wedge}_{ij} = \begin{cases} \dfrac{N_{ij}}{\sum_{k=1}^{N_{ij}} T_{ijk}}, & \text{for } i \neq j, \\ -\sum_{j \neq i} \lambda_{ij}, & \text{for } i = j, \end{cases} \tag{3}$$

where $\lambda^{\wedge}_{ij}$ – the non-negative intensity of state transition from state S_i to $S_{j \neq i}$ (exits S_i to $S_{j \neq i}$; random variable); N_{ij} – number of S_i state observations before state S_j (random variable), equal to the number of transitions $S_i \longrightarrow S_{j \neq i}$; $\{T_{ijk}\}_{1 \leq k \leq N_{ij}}$ – the sequence of sojourn times when the system takes the state S_i and transitions to state S_j in the interval $[t - t_0, t]$.

Since the estimators of the transition intensities have not been developed, the endpoints of the confidence intervals for the intensities $\lambda^{\wedge}_{ij}$ are calculated as the reciprocals of the endpoints of the average times $\overline{T_{ij}}$:

$$\lambda^{\wedge}_{1ij} = 1/\overline{T_{2ij}}; \quad \lambda^{\wedge}_{2ij} = 1/\overline{T_{1ij}} \quad \text{for } \overline{T_{1ij}} < \overline{T_{2ij}}. \tag{4}$$

For any time distribution T_{ij} of the duration of the initial transition state and sample size of at least 100, a normal distribution of the sample mean is assumed based on the Central Limit Theorem, and the endpoints of the central confidence interval are calculated according to formula (5):

$$\begin{aligned} \overline{T_{1ij}} &= \overline{T_{ij}}\left(n_{ij}\right) - u_{\propto} \cdot \frac{S\left(T_{ij}\right)}{\sqrt{n_{ij}}}, \\ \overline{T_{2ij}} &= \overline{T_{ij}}\left(n_{ij}\right) + u_{\propto} \cdot \frac{S\left(T_{ij}\right)}{\sqrt{n_{ij}}}, \end{aligned} \tag{5}$$

where $\overline{T_{1ij}} < \overline{T_{2ij}}$ – ends of the confidence interval for the mean time $\overline{T_{ij}}$ of the initial transition state; $\overline{T_{ij}}\left(n_{ij}\right) = \overline{T_{ij}}$ – mean duration of the initial transition state from the test; $S(T_{ij})$ – unbiased standard deviation of the empirical time distribution T_{ij} of the initial transition state; n_{ij} – number of state S_i to state S_j transitions during the test; α – risk of incorrect estimation based on an unusual random sample; u_α – bilateral critical value of order α of decomposition $N(0;1)$; $P = 1 - \alpha$: assumed confidence level.

For known time distributions T_{ij} and a small number of tests, the endpoints of confidence intervals are calculated according to individual formulas. For example, for the exponential time distribution T_{ij}, the asymmetrical confidence interval for the mean time $\overline{T_{ij}}$ is defined by formulas [60]:

$$\begin{aligned} \overline{T_{1ij}} &= \overline{T_{ij,\beta}} = \frac{2\sum_{k=1}^{n_{ij}} T_{ijk}}{\chi^2_{\gamma,2n_{ij}}} = \frac{2n_{ij}\overline{T_{ij}}}{\chi^2_{\gamma,2n_{ij}}}, \\ \overline{T_{2ij}} &= \overline{T_{ij,\gamma}} = \frac{2\sum_{k=1}^{n_{ij}} T_{ijk}}{\chi^2_{\beta,2n_{ij}}} = \frac{2n_{ij}\overline{T_{ij}}}{\chi^2_{\beta,2n_{ij}}}, \end{aligned} \tag{6}$$

where $k = 1$; n_{ij} – number of time T_{ij} observations during the test; $\chi^2_{\beta,2nij}$; $\chi^2_{\gamma,2nij}$ – critical values of rows β and γ and distribution χ^2 with $2n_{ij}$ degrees of freedom. Other markings as above.

The confidence intervals for the parameters of equations (1.1)–(1.5) are calculated according to (6) for the transition intensities of $\lambda^{\wedge}_{ij}$, except for the parameters $\lambda^{\wedge}_{ii}$, for which the endpoints of the confidence interval are calculated according to the following formulas (7), based on the definition (5) of diagonal transition intensities as

$$\begin{aligned} \lambda^{\wedge}_{1ii} &= -\sum_{j \neq i} \lambda^{\wedge}_{2ij}, \\ \lambda^{\wedge}_{2ii} &= -\sum_{j \neq i} \lambda^{\wedge}_{1ij}, \\ \lambda^{\wedge}_{1ii} &< \lambda^{\wedge}_{2ii}. \end{aligned} \tag{7}$$

The confidence intervals for the parameters a_{ij} of models (2.1)–(2.5) are calculated according to definition (8)

$$P(a_{1ij} < a_{ij} < a_{2ij}) = 1 - \alpha, \tag{8}$$

where α is the given confidence level; a_{1ij}; a_{2ij} ends calculations can be done by many programs for analysis, like Gretl, for example.

Stage 4 of the procedure for verifying the Chapman-Kolmogorov assumption is to test the coefficient of determination and the significance of the model parameter estimates a_{ij} (2), and conduct a comparative analysis of the relationship between confidence intervals for parameters a_{ij} of models (2.1)–(2.5) and λ_{ij} Chapman-Kolmogorov assumption parameters (1.1)–(1.5). The results of the analysis may be as follows:

- If all estimates of the structural parameters of model (2) for the difference quotient $r_j(t)$ prove to be insignificant, the Chapman-Kolmogorov assumption is wrong.
- If all estimates of the structural parameters of model (3) for the difference quotient $r_i(t)$ prove to be significant and the model coefficient of determination is satisfactory (minimum 0.5 or higher, for example 0.81), the Chapman-Kolmogorov assumption for the difference quotient $r_i(t)$ proves to be generally formally correct (*reductio ad absurdum* did not occur). However, additional contextual validation is required, based on investigating the confidence intervals for the intensity λ_{ij} of equations (1.1)–(1.5) and parameters a_{ij} in the regression model (2).
- If all pairs of confidence intervals have a nonempty common part, the Chapman-Kolmogorov assumption for the difference quotient $r_j(t)$ proves to be correct, both formally and contextually. This is described in (9)

$$(a_{1ij};\ a_{2ij}) \cap (\lambda^{\wedge}_{1ij};\ \lambda^{\wedge}_{2ij}) \neq \emptyset. \tag{9}$$

In other cases, the Chapman-Kolmogorov assumption for the difference quotient $r_j(t)$ is formally and/or contextually correct only for certain structural parameters. This assumption may also prove to be formally incorrect but contextually correct if the coefficient of determination of the regression model (2) is low (< 0.5) or the data sample is too small.

Formally, the correct estimation of parameters of multiple linear regression models assumes the fulfilment of the condition of the regressor independence (independent variables) and the selection of an estimation method (considering the parasitic factors) that results in a normal distribution of model residuals.

In the case of a stochastic process with a deterministic component that occurs in real processes of operation, the condition of independence of regressors in model (2) may not be satisfactorily met because the regressors may be interdependent. The manifestation of interdependency was the strong correlations of regressors that occurred for the operation process under study.

In such a case, the estimation of regression model (2) using the OLS method is only a specific, not necessarily optimal fitting of the structural parameters to the empirical data [61, 62], the quality of which is determined by the coefficient of determination, fitting errors, the significance of the slope coefficient, and the degree of compliance of the distribution of residuals with the normal distribution.

The improvement of model (2) practiced in econometrics by eliminating parasitic factors is not possible in the case of the verification procedure, since the verified Chapman-Kolmogorov assumption is constant and not subject to modification. One can only perform the estimation of models (2.1)–(2.5) with correction of heteroscedasticity relative to the observation number and compare the estimation results of models (2.1)–(2.5) without and with heteroskedasticity correction.

The proposed procedure for verifying the Chapman-Kolmogorov assumption had to be validated by examining the regression nonlinearity of simple multiple regression components of (2) and regression nonlinearity tests from (2), and performing the verification procedure for a simulated stochastic process without deterministic deformations of the stochastic matrix and the intensity of transitions observed in the real vehicle operation process.

For this purpose, based on its stochastic matrix and average state durations, random coupled phase trajectories in discrete and continuous time were generated using the programs *Mathematica* and *Gretl*.

3. DESCRIPTION OF THE EXAMINED PROCESS OF OPERATION AND THE RESULTS OBTAINED

This study examined the operational process of nine military vehicles in 2019 and 2020. Each object could be in one of five distinct operational states. The states were: S_1 – performing the task; S_2 – refueling; S_3 – garage parking; S_4 – ongoing maintenance; S_5 – periodic maintenance and repairs. States S_1 and S_4 occurred with equal frequency but were randomly interspersed with states S_2 and S_3, and therefore could not be integrated. Based on operational records, statistical databases were created and then processed into collective phase trajectories for sets of states and their durations. No significant differences were found in the distribution of state frequencies and their durations over the research period.

However, significant seasonality in the intensity of vehicle operation was observed across hours of the day, days of the month, days of the week, months, and quarters of the year. Nonetheless, the multi-seasonality affected only the number of state observations, not the frequency or duration of the states. This was advantageous for the present study, provided that the Markov model matrix was estimated over a sufficiently lengthy period, at least one year. Autocorrelation of states and deviations of their durations from exponential distributions were also examined; however, their relevance to the goals of this study was not established. The year 2019 was used as the period of expired forecasts and model matrix estimation, while 2020 was adopted for *ex ante* forecasts.

3.1. Results of the estimations of the probabilities and intensity of the process under study transitions

Table 1 presents the matrices of interstate transition frequencies and intensities for the vehicle sample under study.

The process under study did not show significant autocorrelation of states, it had a nonsingular stochastic matrix $[w_{ij}]$ and satisfied the conditions of stationarity and ergodicity for the set of states over at least six-month periods, whereas the state probabilities p_j in 2019 were very close to equilibrium with the limit values $p_{j\infty}$ in discrete time (Table 2).

Table 1
Matrices of frequencies w_{ij} and intensities of transitions $\lambda^{\wedge}_{ij}$ (in 1/day) of the process in 2019

w_{ij}	S_1	S_2	S_3	S_4	S_5
S_1	0	0.1310	0	0.8690	0
S_2	0	0	0.7317	0.2683	0
S_3	0.6078	0.2235	0	0	0.1686
S_4	0	0	1	0	0
S_5	0.3023	0.0698	0.6279	0	0

$\lambda^{\wedge}_{ij}$	S_1	S_2	S_3	S_4	S_5
S_1	–4.3236	1.8859	0	2.4377	0
S_2	0	–962.25	518.97	443.28	0
S_3	0.09901	0.32910	–0.46429	0	0.03618
S_4	0	0	144	–144	0
S_5	0.18303	0.32251	0.96214	0	–1.46767

Table 2
Frequencies w_j %, limiting probabilities $p_{j\infty}$ % (calculated with Mathematica program), and their percentage deviations δ % from frequencies and quotients p_j/w_j

Discrete time, 2019					
Parameter	S_1	S_2	S_3	S_4	S_5
w_j %	23.172	11.310	36.414	23.172	5.931
$p_{j\infty}$ %	23.464	11.453	35.615	23.464	6.006
δ %	1.26	1.264	-2.194	1.26	1.26
p_j/w_j	1.013	1.013	0.978	1.013	1.013

Continuous time, 2019					
Parameter	S_1	S_2	S_3	S_4	S_5
w_j %	2.178	0.0050	94.48	0.0355	3.2996
$p_{j\infty}$ %	2.279	0.0378	95.18	0.1550	2.3462
δ %	4.621	656	0.7414	336.3	–28.89
p_j/w_j	1.046	7.516	1.007	4.363	0.711

The assumption of exponentially distributed state durations was also satisfactorily met.

The frequencies of states in discrete and continuous time differed significantly, which is common for operational processes. In continuous time, the vehicle was parked S_3 due to the low intensity of operation. Several hundred percent variations of limiting probabilities $p_{j\infty}$ % of some states (S_2 and S_4) relative to their frequency in continuous time, were found (Table 2). In the context of forecasting a modest increase in the probability of task performance state S_1 (4.6%), the results of the Chapman-Kolmogorov equations (1.1)–(1.7) suggested several hundred percent increases in the probability of refueling S_2 and ongoing maintenance S_4, and a significant decrease in the probability of periodic maintenance and repairs S_5. These forecasts are unreliable from the standpoint of vehicle operation, which in turn makes the examined process a suitable object for verifying the Chapman-Kolmogorov assumptions (1.1)–(1.5). To perform the verification, the width $t - t_0$ and the time intervals $\langle t_0; t\rangle$ had to be determined so that the estimation of the $(\mathrm{d}/\mathrm{d}t)p_i$ probability derivatives using difference quotients according to formula (2) were resistant to missing state observations in the interval $\langle t_0; t\rangle$. For this purpose, the determination of process state probabilities in discrete time was examined. Satisfactory accuracy of each probability estimate, with a deviation of less than 5%, occurred after 15–20 steps of the process, while stabilization is clearly visible after 30 steps. This phenomenon is illustrated by an example graph of the function $p_5(t)$ for state S_5 in Fig. 2. This determines the starting time t_0 and the width t-t_0 of the time interval for estimating derivatives according to (2) as a minimum of 15–20 steps.

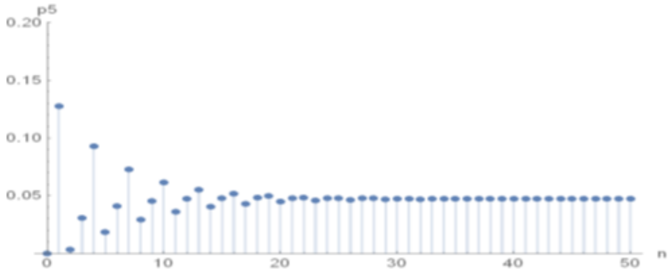

Fig. 2. Evolution of $p_5(t)$ repairs and periodic maintenance probability after n steps of the vehicle operation process from the garage state S_3 in step $n = 0$ according to Mathematica

3.2. Results of estimating confidence intervals for transition intensities

The tested process showed a very large variation in the number of transitions n_{ij} between states (Table 3), and, for this reason, the asymptotic formula (5) for the endpoints of the confidence interval of mean times could not be used for all transitions $\overline{T_{ij}}$ (duration of transition states $S_i \to S_j$). Formula (6) was used for transition numbers 10–99. For the occasional transition $S_5 \to S_2$, the endpoints of the confidence interval were calculated according to (4), taking the minimum and maximum values of observations T_{52} from the test.

Endpoints of confidence intervals for diagonal intensities λ_{ii} were calculated as the negative sums of the endpoints of the confidence intervals of off-diagonal row "i" elements of the intensity matrix

$$\begin{aligned} \lambda^{\wedge}_{1ii} &= -\sum_{j\neq i} \lambda^{\wedge}_{2ij}\,, \\ \lambda^{\wedge}_{2ii} &= -\sum_{j\neq i} \lambda^{\wedge}_{1ij}\,, \\ \lambda^{\wedge}_{1ii} &< \lambda^{\wedge}_{2ii}\,. \end{aligned} \tag{10}$$

Table 3

Results of estimating the endpoints of confidence intervals (0.95) of mean times $\overline{T_{ij}}$ (min) of the initial transition states and transition intensities $\lambda_{ij}^{\wedge}$ (1/min) in 2019

S_i	S_j	n_{ij}	$\overline{T_{ij}}$	$\overline{T_{1ij}}$	$\overline{T_{2ij}}$	$\lambda_{ij}^{\wedge}$	$\lambda_{1ij}^{\wedge}$	$\lambda_{2ij}^{\wedge}$
1	2	22	763.5697	523.3069	1218.408	0.00131	0.000821	0.001911
1	4	146	590.7192	355.1579	826.2805	0.001693	0.00121	0.002816
2	3	60	2.774722	2.187528	3.636093	0.360396	0.27502	0.457137
2	4	22	3.248485	2.226325	5.183521	0.307836	0.192919	0.449171
3	1	155	14544.02	9520.069	19567.97	6.88E-05	5.11E-05	0.000105
3	2	57	4375.526	3429.631	5777.109	0.000229	0.000173	0.000292
3	5	43	39803.42	30771.53	54999.51	2.51E-05	1.82E-05	3.25E-05
4	3	168	10	9.984878	10.01512	0.1	0.099849	0.100151
5	1	13	7867.769	4879.45	14776.32	0.000127	6.77E-05	0.000205
5	2	3	4465	4700	7220	0.000224	0.000139	0.000213
5	3	27	1496.667	1060.741	2271.096	0.000668	0.00044	0.000943

Table 4

Confidence intervals (0.95) of diagonal intensities λ_{ii} of transitions (1/min) in 2019

	λ_{11}	λ_{22}	λ_{33}	λ_{44}	λ_{55}
λ_{ii} from the test	–0.003002	–0.668232	–0.000322	–0.100000	–0.001019
λ_{1ii}	–0.004727	–0.906308	–0.000429	–0.100151	–0.001360
λ_{2ii}	–0.002031	–0.467940	–0.000242	–0.099849	–0.000646

3.3. Estimation results of linear regression models for difference ratios r_j $(p_{i\neq j})$

The probability estimators $p_i(t)$ of the Chapman-Kolmogorov equations are the cumulative frequencies $w_j(t)$ of the regression model equations (2), calculated in continuous time according to formula (11)

$$w_i(t) = \frac{\sum\limits_{j=1}^{n_i(t)} T_{ij}}{\sum\limits_{i}\sum\limits_{j=1}^{n_i(t)} T_{ij}}, \tag{11}$$

where $\sum\limits_{j=1}^{n_i(t)} T_{ij}$ – sum of the durations of state no. i to t moment; $\sum\limits_{i}\sum\limits_{j=1}^{n_i(t)} T_{ij}$ – of all state duration times to t moment; j – number of the state observation no. i in the data sample; $n_i(t)$ – quantity of the state observation no. i to t moment.

Unlike discrete time, for continuous time the sums indicated in equation (11) refer to the duration of states, not the number of states. Cumulative frequency plots W_{t1}–W_{t5} of the process under study, as tested in Fig. 3 and Fig. 4, illustrate the obstacles in the verification of hypotheses about continuous-in-time process state probability distributions. Obstacle 1 – the discontinuity of the distribution of observations – hinders the verification of point hypotheses about the derivatives of probabilities at time t. Only interval estimation of derivatives as difference quotients according to formula (2) is possible. Obstacle 2 is large, random frequency fluctuations $W_{tj}(t)$ occurring in the initial period of ex ante forecasts (Fig. 4). For the tested process, the fluctuations died out after 15–20 steps of the discrete-time process (Fig. 2), which corresponded to approximately 40 000 minutes (about 28 days) of continuous time. This period was a prognostically unavailable period with one data sample. It would be available for data on hundreds of parallel processes that were not implemented in the operation of military vehicles.

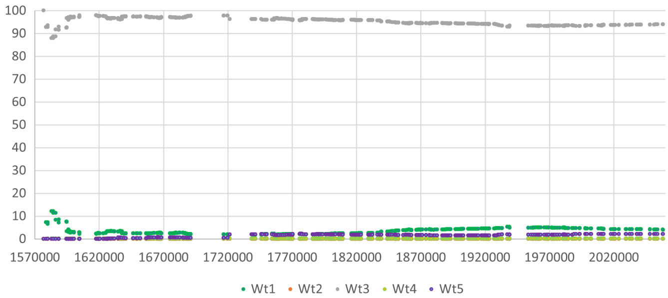

Fig. 3. Cumulative percentage frequencies $W_{t1}(t)$–$W_{t5}(t)$ states of the process under study in 2020

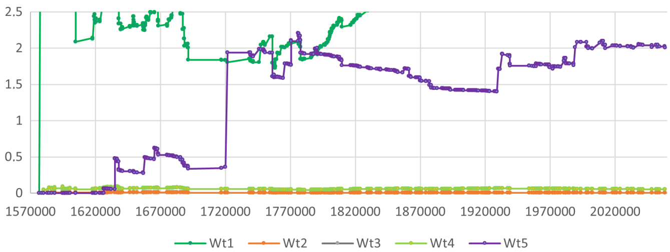

Fig. 4. Cumulative states percentage frequencies $W_{t1}(t)$–$W_{t5}(t)$ of the process under study without S_3 state in 2020

Time *vol* (in minutes) was measured from the beginning of 2017. The dominance of garage time is visible as $W_{t3}(t)$.

Obstacle 2 described above limits forecasting to the time for which fluctuations in cumulative state frequencies remain smaller than the accepted forecast error. Obstacle 3 consisted in interruptions in the operation of vehicles when they were parked (constant frequencies in Fig. 3 and Fig. 4). These interruptions were both seasonal and organizational breaks that altered the environment of the process under study. For example, a fivefold jump in frequency W_{t5} around approximately 1 720 000 minutes (Fig. 4) corresponds to the end of a waiting period of about 20 days for spare parts. Environmental factors should also be attributed to the weekly breaks in vehicle operation near 1 950 000 minutes, frequent operational pauses in quarter 4 (for $t > 2\,000\,000$ min), and the complete operational downtime in December 2020 ($t > 2\,060\,000$ min), when reports are generated and many public holidays occur. These obstacles

Table 5

The confidence intervals (0.95) for the parameters of equations (1.1)–(1.5) throughout 2019 and the regression model equations (2) for the derived probabilities in the period from 10/06/2020 to 04/09/2020

According to the estimation of the intensity of transitions				Common part
Parameter	Sample	min	max	
λ_{11}	–0.003	–0.0047	–0.002	$\emptyset$
λ_{12}	0.00131	0.00082	0.00191	$\neq \emptyset$
λ_{14}	0.00169	0.00121	0.00282	$\neq \emptyset$
λ_{22}	–0.6682	–0.9063	–0.4679	$\emptyset$
λ_{23}	0.3604	0.27502	0.45714	$\emptyset$
λ_{24}	0.30784	0.19292	0.44917	$\emptyset$
λ_{31}	6.9E-05	5.1E-05	0.00011	$\emptyset$
λ_{32}	0.00023	0.00017	0.00029	$\emptyset$
λ_{33}	–0.0003	–0.0004	–0.0002	$\emptyset$
λ_{35}	2.5E-05	1.8E-05	3.2E-05	$\emptyset$
λ_{43}	0.1	0.0998	0.1002	$\emptyset$
λ_{44}	–0.1	–0.1002	–0.0998	$\emptyset$
λ_{51}	0.00013	6.8E-05	0.000205	$\emptyset$
λ_{52}	0.00022	0.00014	0.00068	$\neq \emptyset$
λ_{53}	0.00067	0.00044	0.000943	$\emptyset$
λ_{55}	–0.001	–0.0014	–0.0006	$\emptyset$

According to the estimation of regression models (2.1)–(2.5)					
Parameter	Sample	min	max	Essence	Model
a_{11}	–1.1E-05	–1.6E-05	–6.7E-06	*YES*	(3.1) R^2 0.6190
a_{12}	–0.00018	–0.00919	0.008839	*NO*	
a_{14}	0.001108	–4.6E-05	0.002262	*YES*	
a_{22}	4.95E-05	3.54E-05	6.36E-05	*YES*	(3.2) R^2 0.1880
a_{23}	-3.6E-09	–4.6E-09	–2.6E-09	*YES*	
a_{24}	3.24E-07	–2.4E-07	8.9E-07	*NO*	
a_{31}	-3.3E-06	–7.1E-06	5.09E-07	*YES*	(3.3) R^2 0.4143
a_{32}	–0.06174	–0.08702	–0.03645	*YES*	
a_{33}	5.49E-06	3.87E-06	7.11E-06	*YES*	
a_{35}	–7.84E-05	–9.7E-05	–6E-05	*YES*	
a_{43}	–1.62E-08	–2.2E-08	–1.1E-08	*YES*	(3.4) 0.1315
a_{44}	2.61E-05	1.74E-05	3.49E-05	*YES*	
a_{51}	9.84E-06	7.94E-06	1.18E-05	*YES*	(3.5) R^2 0.4860
a_{52}	–0.007	–0.01966	0.005667	*NO*	
a_{53}	–8.4E-07	–1.6E-06	–2.5E-08	*YES*	
a_{55}	5.35E-05	4.44E-05	6.26E-05	*YES*	

Explanations for symbols:
Parameters of equations (1.1)–(1.5) according to the estimation of the intensity of transitions (Table 3, Table 4) in 2019. Parameters of equations (2.1)–(2.5) according to the Gretl program for 231 data from the period from 10/06/2020 to 04/09/2020.
Sample – the value of the sample parameter;
min, max – endpoints of confidence intervals.
R^2 – coefficient of determination.
Essence – significance of the estimate > 0.95 *YES*/*NO*.
Common Part – the common part of the confidence intervals for λ_{ij} and a_{ij}: $\emptyset$ – empty set, $\neq \emptyset$ nonempty set.

limited the verification of *ex ante* forecasts of the Chapman-Kolmogorov system of equations for 2020 to the period from 10/06/2020 to 04/09/2020 and restricted the forecasting horizon. Additionally, there was also obstacle 3, related to the verification of models of military vehicle operation processes in peacetime: a remarkably low intensity of operation and the overwhelming dominance of the garage state S_3 frequency W_{t3} (Fig. 3). This reduced the sample size and increased estimation errors.

OLS estimations were performed using moving variables of model (2), as this approach utilizes the data more effectively than reducing process fluctuations by applying a moving average smoothing method. This was confirmed by positive results from normality tests of the residuals of models (2), which fell within the three-sigma range for four out of five model (2) estimates. An example histogram of the residuals of model (2.5) for the difference quotient $r_5(t)$ of the state probability S_5, with a positive χ^2 test result, is shown in Fig. 5.

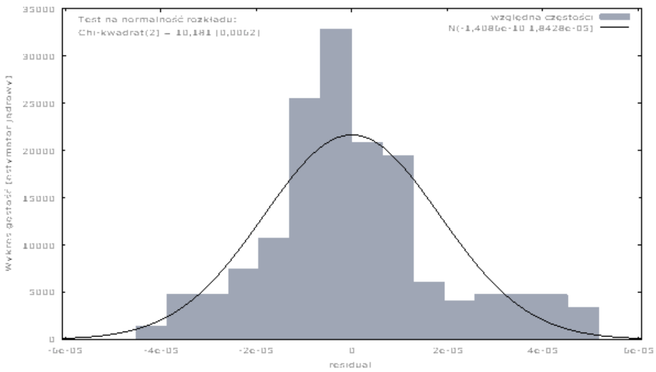

Fig. 5. Histogram of residuals and χ^2 normality test result of the difference quotient $r_5(t)$ model (2.5)

No other post-estimation tests were performed because the purpose of estimating models (2) was solely to fit the parameters to the data, and it was not possible to evaluate the models beforehand. According to the estimation results presented in Table 5, the Chapman-Kolmogorov assumption proved to be statistically inconsistent with the sample data for 13 out of 16 parameters of equations (1.1)–(1.5) and the regression models (2). Only for three parameter pairs (λ_{ij} and a_{ij}) were there nonempty intersections of confidence intervals, and only one model (2.1) for the difference quotient $r_1(t)$ had a coefficient of determination $R^2 = 0.619$, indicating that the model variance predominated over random variance. Model (2.5) nearly met this condition, also showing a nonempty common part of the confidence intervals for the parameter pairs (λ_{52}; a_{52}). The high statistical significance of most coefficient a_{ij} in regression models (2) (≥ 0.95) despite small determination coefficients was a consequence of the large sample size (231) and the compensation for variable fluctuations $w_i(t)$ in model (2), which are not strictly linearly independent due to the normalization conditions (1.6) and (11). The negative verification of the applicability of the Chapman-Kolmogorov assumption for the examined operation process required a fourfold validation of the proposed verification procedure:

- Validation *of ex ante* forecasts for the average probabilities of states during the *ex ante* forecast period.
- Validation of *ex ante* forecast for the endpoints of the confidence intervals of transition intensities (Tables 3, 4).
- Validation of the verification procedure results with tests of nonlinearity for models (3) and simple regressions.
- Validation of the proposed verification procedure using a simulated Markov process without seasonal and deterministic deformations of the vehicle operation process matrix under study.

4. CHAPMAN-KOLMOGOROV ASSUMPTIONS VERIFICATION PROCEDURE VALIDATION RESULTS

4.1. Average rates and ex ante forecasts for the average state probabilities

For continuous time (in minutes) from the beginning of 2020, the *ex ante* forecast period from 10/06/2020 to 03/09/2020 spanned the range [241 705; 356 390] minutes. According to the Chapman-Kolmogorov system with estimates based on 2019 data, the examined process stabilized with deviations from the limiting values $p_{j\infty}$ below 0.01% after 4000 minutes ($p_1(t)$ in Fig. 6). During *ex ante* forecast period, the state probabilities differed from their limiting values by a maximum of 0.00000001%, meaning that the *ex ante* forecasts coincided with the limiting forecasts.

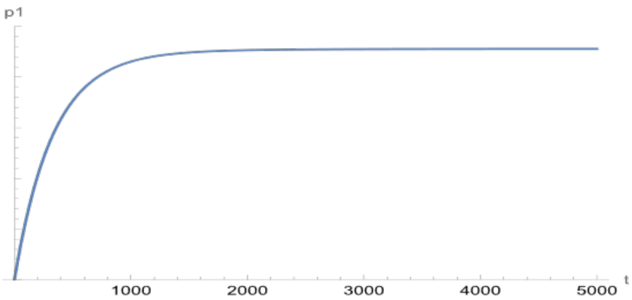

Fig. 6. Evolution of S_1 state probability over time from 0 to 5000 minutes, calculated using the Mathematica program

Only the forecast for the probability of state S_3 (garage) met the forecast deviation condition δ% within ±20%, which is commonly accepted in academic publications, and was within the confidence interval for points, but not for means. Predictions for the probabilities of state S_1 (task execution) and S_5 (repair and periodic maintenance) were inaccurate and exceeded the confidence intervals for both means and points, but they met the deviation condition δ% within ±50%, which is accepted in econometrics. Forecasts for short-term states S_2 (refueling) and S_4 (ongoing maintenance) were nonsensical, with deviations of 483% and 157%, respectively.

4.2. Forecasts validation results for the endpoints of transition intensity confidence intervals

Forecasts of the limiting state probabilities $p_{j\infty\%}$ based on the endpoints of the confidence intervals for transition intensities (lower and upper estimates in Table 7) proved to be similar to the previously discussed central estimates (Table 2), and confirmed the drastic failure to meet the Chapman-Kolmogorov assumption for the probabilities of states S_2 and S_4.

Based on the data presented in Table 7, it can be assumed that the errors in the limiting forecasts of the Chapman-Kolmogorov system of equations are much more sensitive to deviations from the Chapman-Kolmogorov assumption than to errors in the estimation of transition intensities.

Confirming this assumption, however, would require complex sensitivity analyses of the forecasts.

Table 6

Mean condition frequencies, confidence intervals (0.95), and *ex ante* forecasts for the average probabilities of states of the process under study in the period from 10/06/2020 to 04/09/2020

Parameter	S_1	S_2	S_3	S_4	S_5
Condition frequency w_{ti}% mean	3.777105	0.006487	94.58	0.060244	1.576165
Probability p_{ti}% mean	2.27898	0.037806	95.182	0.154961	2.34622
Deviation δ% (p_{ti}% mean od w_{ti}% mean)	−39.6633	482.8239	0.636498	157.2238	48.85626
SD% (w_{ti}%)	0.703123	0.000108	0.614934	0.002544	0.143637
SD% (w_{ti}% mean)	0.046262	7.07E-06	0.04046	0.000167	0.009451
Endpoint of 1 confidence interval (0.95) w_{ti}%	2.398984	0.006276	93.37473	0.055258	1.294637
Endpoint of 2 confidence interval (0.95) w_{ti}%	5.155225	0.006697	95.78527	0.065229	1.857693
Endpoint of 1 confidence interval (0.95) w_{ti}% mean	3.686431	0.006473	94.5007	0.059916	1.557642
Endpoint of 2 confidence interval (0.95) w_{ti}% mean	3.867778	0.006501	94.6593	0.060572	1.594688
Number of observations n_i	231	231	231	231	231

Explanations for symbols:
The letter t in a parameter label indicates that the value is in continuous time, not discrete time. Standard deviations (SD) and confidence intervals are provided for both point values and mean values over the *ex ante* forecast period because the Chapman-Kolmogorov system predictions remained constant to ten decimal places throughout this period.

Table 7
Limiting forecasts for the endpoints of the transition intensity confidence intervals

Parameter	S_1	S_2	S_3	S_4	S_5	Type of estimation
Condition frequency w_{tj}% mean	3.777105	0.006487	94.58	0.060244	1.576165	
Probability $p_j\infty$%	2.27898	0.037806	95.182	0.154961	2.34622	Central estimation
Deviation D% ($p_j\infty$% od w_{tj}% mean)	–39.6633	482.8239	0.636498	157.2238	48.85626	
Probability $p_j\infty$%	2.51312	0.028431	95.6341	0.118114	1.70623	Lower estimation
Deviation D% ($p_j\infty$% od w_{tj}% mean)	–33.4644	338.2968	1.114506	96.06054	8.252004	
Probability $p_j\infty$%	2.21545	0.035849	95.2503	0.223066	2.27529	Upper estimation
Deviation D% ($p_j\infty$% od w_{tj}% mean)	–41.3453	452.6531	0.708712	270.2731	44.3561	

4.3. Results of testing the nonlinearity of multiple and simple regression models

Nonlinearity tests of the Lagrange multiplier (LM) performed by the Gretl program showed the significance of the coefficients of squares and logarithms of the variables for all five multiple regression models (2.1)–(2.5), as shown in Table 5. An example of the test result for model (2.1) is presented in Table 8.

Table 8
The auxiliary regression equation for the nonlinearity test (squared variables). OLS estimation, observations used 471–701 ($n = 231$); dependent variable (Y): uhat

	Coefficient	Standard error	t-Student	Value of p
w_{t1}	9.97150E-06	1.62599E-05	0.6133	0.5403
w_{t2}	−0.706905	0.181267	−3.900	0.0001***
w_{t4}	0.0656618	0.0195452	3.359	0.0009***
sq_w_{t1}	−1.31285E-06	2.23952E-06	−0.5862	0.5583
sq_w_{t2}	59.4515	13.8490	4.293	2.62E-05***
sq_w_{t4}	−0.520685	0.158828	−3.278	0.0012***

Coefficient of determination. R-square = 0.301980.
Test statistics: TR^2 = 69.7575, with a value of
p = P(Chi-square(3) > 69.7575) = 4.8105e-015
***significance at 1%; **significance at 5%; *significance at 10%.

In terms of verifying the Chapman-Kolmogorov assumption, the results of the LM tests are only guidelines for testing the non-linearity of simple regressions, as they consider only two specific types of nonlinearity and, due to the doubling of the number of regressors, may distort the results of significance tests of structural parameters and the value of the model coefficient of determination.

An example is model (2.1), where parameter a_{12} of the linear model (Table 5) proved to be significant according to the LM test for the model with squared variables, while the coefficient of determination of the quadratic regression model (0.302) was more than twice as small as that of the linear regression model (0.619 – Table 5).

A more dependable and, at the same time, demonstrative validation of the proposed procedure is possible by comparing the coefficients of determination of simple linear and non-linear regressions for the dependent variables of equations (2.1)–(2.5) in relation to individual regressors.

The comparison results are presented in Table 9. The corresponding plots show the dispersion of dependent variables against regressors. Polynomial regression was tested as a more general method than quadratic and logarithmic regression.

Validation of the proposed procedure using simple and multiple regression tests demonstrated the advantage of the condition involving the existence of a common part of the confidence intervals for the transition intensity estimates and the slope coefficients of multiple regression models, as a criterion for the satisfactory local fulfilment of the Chapman-Kolmogorov assumption, over the ambiguous relationships indicated by the determination coefficients of linear and polynomial simple and multiple regression models.

This ambiguity results from the strong definitional sensitivity of the determination coefficient to small but significant deviations of data from mean values. This is illustrated in Fig. 7.

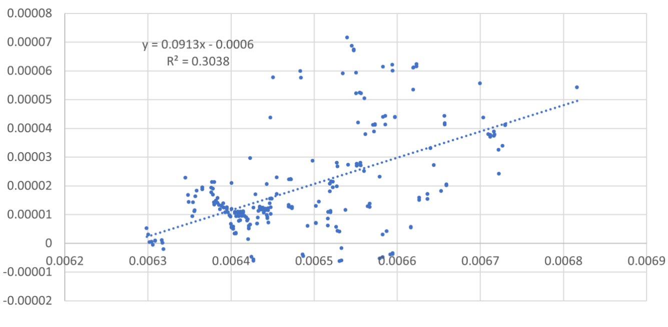

Fig. 7. Difference quotient of dispersion plot r_1 (%/min) and frequency w_{t2} (%) with the linear regression line

A textbook dispersion plot for linear regression, shown in Fig. 7, illustrates the concentration of 90% of observation pairs close to the simple linear regression line, and approximately 10% of outlier pairs, which lowered the coefficient of determination of the linear regression to an unsatisfactory value of 0.3038 < 0.5.

Table 9

List of values of the coefficient of determination of simple linear and nonlinear polynomial regressions for dependent variables in relation to the regressors of models (2.1)–(2.5)

Model	Dependent variable	Regressor	R^2 linear	R^2 polynomial	Degree	Validation	Common part of confidence intervals
(3.1) R^2 0.6190	r_1	w_{t1}	0.087	0.618	5	*YES*	$\emptyset$
	r_1	w_{t2}	0.3038	0.3334	6	*YES*	$\neq \emptyset$
	r_1	w_{t4}	0.0194	0.1778	5	*YES*	$\neq \emptyset$
(3.2) R^2 0.1880	r_2	w_{t2}	0.1585	0.1963	6	*YES*	$\emptyset$
	r_2	w_{t3}	0.014	0.1725	5	*YES*	$\emptyset$
	r_2	w_{t4}	0.0128	0.0963	6	*YES*	$\emptyset$
(3.3) R^2 0.4143	r_3	w_{t1}	0.031	0.5179	5	*YES*	$\emptyset$
	r_3	w_{t2}	0.2428	0.2642	5	*YES*	$\emptyset$
	r_3	w_{t3}	0.0279	0.4666	6	*YES*	$\emptyset$
	r_3	w_{t5}	0.1875	0.387	4	*YES*	$\emptyset$
(3.4) R^2 0.1315	r_4	w_{t3}	0.0008	0.251	6	*YES*	$\emptyset$
	r_4	w_{t4}	0.1815	0.2247	5	*YES*	$\emptyset$
(3.5) R^2 0.4860	r_5	w_{t1}	0.0761	0.5407	5	*YES*	$\emptyset$
	r_5	w_{t2}	0.0116	0.0934	6	?	$\neq \emptyset$
	r_5	w_{t3}	0.1493	0.794	3	*YES*	$\emptyset$
	r_5	w_{t5}	0.0865	0.6696	5	*YES*	$\emptyset$

Explanations for symbols:

Degree – The highest degree of polynomial regression that does not significantly increase the value R^2.

Validation – Yes – empty or nonempty common part of confidence intervals as given in Table 5, according to the following validation criteria for the values and relations of the coefficient of determination of simple linear and polynomial regression and the coefficient of determination of multiple regression:

- Nonempty part common to R^2 multiple regression > 0.5 and R^2 polynomial regression < 0.5.
- Empty shared area for R^2 multiple regression < 0.5.

Validation – ? – Too few transitions $n_{52} = 3$ (Table 3) for the existence of a nonempty common part could be considered reliable, which may have been the result of a small sample.

After eliminating the outlier pairs, the coefficient of determination increased to 0.725 > 0.5, confirming the Chapman-Kolmogorov assumption according to the relationship between the confidence intervals.

The dispersion plot in Fig. 8 illustrates the concentration of 80% observation pairs close to the simple linear regression line and approximately 20% of outlier pairs, which lowered the coefficient of determination of the linear regression to an extremely low value of 0.0194.

Once the outlier pairs are eliminated, the coefficient of determination increases to 0.638 > 0.5, confirming the Chapman-Kolmogorov assumption according to the relationship between the confidence intervals.

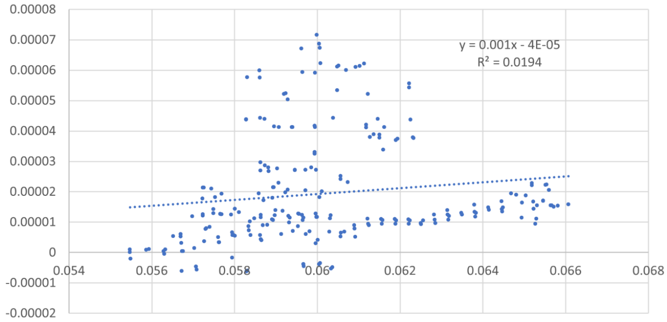

Fig. 8. Difference quotient of dispersion plot r_1 (%/min) and frequency w_{t4} (%) with the linear regression line

The proposed procedure therefore replaces the burdensome elimination of outlier observations that lower the coefficient of determination by filtering them through estimation formulas for the endpoints of confidence intervals at a given confidence level.

In the case of significant nonlinear simple regression (Fig. 9), the proposed procedure detected local nonfulfillment of the Chapman-Kolmogorov assumptions, as indicated by the empty intersection of confidence intervals, despite the satisfactory coefficient of determination of the multiple regression model (2.1) (Table 5). According to the list of determination coefficients for multiple and simple linear and polynomial regression models in Table 9, the Chapman-Kolmogorov assumption was considered not fulfilled if the coefficient of determination of the simple polynomial regression model was greater than 0.5, or greater than 0.3 and at least twice as high as that of the simple linear regression. The assumption was considered fulfilled if the coefficient of determination of the multiple linear regression model exceeded 0.5 and the coefficient of determination of the simple polynomial regression was below 0.5. In other cases, the relationship between the coefficients of determination and the results of nonlinearity tests did not provide a decisive criterion for the degree of fulfilment of the Chapman-Kolmogorov assumptions due to the strong influence of outlier data on the coefficients of determination of regression models (Fig. 8). A criterion less sensitive to outliers was the product of the confidence intervals for transition intensities and the parameters of the regression models, calculated according to the proposed procedure (Table 5).

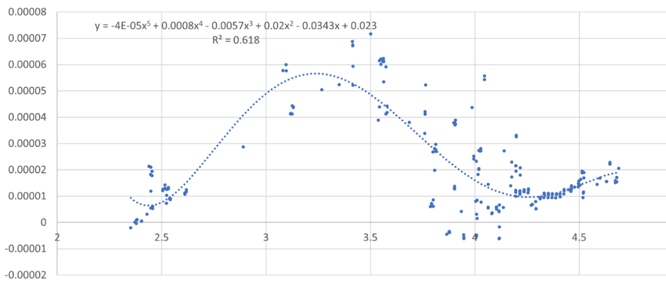

Fig. 9. Difference quotient of dispersion plot r_1 (%/min) and frequency w_{t1} (%) with the power regression line

4.4. Validation results of the proposed verification procedure based on a simulation

To eliminate deterministic deformations of the stochastic matrix and the transition intensities in the real vehicle operation process, random and coupled phase trajectories of 1000 steps in discrete and continuous time were generated using Mathematica and Gretl programs, based on its stochastic matrix and average state durations, with exponential distributions of state durations. The section of steps 1–500 constituted the ex-post period for estimating the matrix of the simulated Markov process. The section of steps 501–687 was the period of process stabilization (Figs. 2–4), and the *ex ante* section of steps 688–1000 was the continuous-time validation period of the proposed procedure. The real and simulated process matrices did not differ significantly, but the phase trajectories simulated in continuous time for the *ex post* and *ex ante* periods were approximately twice as long as the real trajectory due to significant deviations in the "tail" of the real parking time distribution (state S_3) from the simulated exponential distribution.

Compared to the real process (Table 2 and Table 5), the simulated process exhibited the following:

- Significant changes in frequency, limited state probabilities for S_2, S_4, and S_5, and reduced percentage deviations ($\delta\%$) of the limited probabilities of states S_2 and S_4 from their frequency. These effects were a consequence of seasonality, which caused the distributions of the durations of these states to deviate from the simulated exponential distributions.
- Decreases in the coefficient of determination (R^2) for the multiple regression models (2.1)–(2.5) to exceptionally low values, invalidating the Chapman-Kolmogorov assumption for the simulated process. This was confirmed by the absence of overlapping confidence intervals, as shown in Table 5. This result can be considered a positive validation of the proposed procedure using Markov process simulation.

Table 10

Simulation results of the Markov process 1 and determination of model coefficient (2.1)–(2.5)

Determination of the model coefficient of simulation and real models					
Model	(3.1)	(3.2)	(3.3)	(3.4)	(3.5)
R^2 sim	0.2018	0.1103	0.1983	0.0079	0.0815
R^2 real	0.6190	0.1880	0.4143	0.1315	0.4860
$\lvert\lambda_{ii}\rvert$	0.0029	0.6059	0.0002	0.1035	0.0010
$\delta\%$	–11.68	251.06	-0.529	149.94	18.10
Continuous time. Simulation. Forecast periods *ex ante*					
Parameter	S_1	S_2	S_3	S_4	S_5
$w_{ti}\%$	2.338	0.0054	93.72	0.0329	3.9023
$p_i\infty\%$	2.065	0.0191	93.23	0.0823	4.6084
$\delta\%$	–11.679	251.1	–0.5294	149.9	18.10
p_i/w_i	0.883	3.511	0.995	2.499	1.181

Designations as in Table 5, Table 6, Table 7, and Table 9.

The data in the lower left rows of Table 10 suggest an increase in deviations corresponding to the magnitude of the diagonal intensities of limiting transitions ($|\delta\%|$) between Chapman-Kolmogorov forecasts and observed frequencies. However, this could be a result of the dominance of the frequency of state 3. To verify this assumption, a Markov process simulation was performed using the matrix structure of the observed operation process, with higher transition probabilities (1/3, 1/2, and 1) and constant nondiagonal elements of the transition intensity matrix $\lambda_{ij\neq i} = 1$. The results are presented in Table 11.

Table 11

Simulation results of the Markov process 2 and determination of the model coefficient (2.1)–(2.5)

Probabilities p_{ij} and ‖ transition intensity matrix λ_{ij}					
State	S_1	S_2	S_3	S_4	S_5
S_1	0‖ – 2	1/2‖1	0‖0	1/2‖1	0‖0
S_2	0‖0	0‖ – 2	1/2‖1	1/2‖1	0‖0
S_3	1/3‖1	1/3‖1	0‖ – 3	0‖0	1/3‖1
S_4	0‖0	0‖0	1‖	0‖ – 1	0‖0
S_5	1/3‖1	1/3‖1	1/3‖1	0‖0	0‖ – 3
Parameter	S_1	S_2	S_3	S_4	S_5
$w_{ti}\%$	17.6	22.8	34.4	16.0	9.2
$p_i\infty\%$	14.29	21.43	21.43	35.71	7.14
$\delta\%$	−18.83	−6.02	−37.71	123.21	−22.36
$\lvert\lambda_{ii}\rvert$	2	2	3	1	3
Model	(3.1)	(3.2)	(3.3)	(3.4)	(3.5)
R^2 sim	0.1633	0.2300	0.4173	0.3547	0.2347

Designations as in Table 5, Table 6, Table 7, and Table 10.
The matrix element before the ‖ sign is the probability p_{ij}.

A nonmonotonic relationship between $|\delta\%|$ and $|\lambda_{ii}|$ was found, indicating that the relationship observed in Table 10 is specific to the operational process under study and cannot be extrapolated to other Markov processes.

5. SUMMARY

The research presented in this publication aimed to develop a method for verifying the applicability of the Chapman-Kolmogorov assumption for users and researchers of operational processes, which is as simple as possible and feasible using commonly available software. The authors are aware of the preliminary nature of the research and the need for its further development in the future.

According to the cited validation results, the proposed verification procedure for the Chapman-Kolmogorov assumption is equivalent to a set of analyses of the coefficient of determination and the nonlinearity of multiple, simple linear, and nonlinear regression models. By reducing the impact of outlier data, it is reliable and simpler to implement, without requiring database filtering. The results of the Chapman-Kolmogorov assumption verification using the proposed procedure are presented in a compact form, with an indication of the parameters that do not satisfy the assumption. For the operational process under study and two simulations of the Markov process, the Chapman-Kolmogorov assumption was therefore found to be 80–100% incorrect and appeared to serve more as a theoretical affirmation of the surrounding reality.

Unsatisfactory coefficients of determination with high statistical significance are a manifestation of the well-known *large sample size effect*, referred to in the life and medical sciences as the collision of effect size (Cohen's) with test significance. This effect means that, with a sufficiently large sample, even an insignificant parameter can appear statistically significant in a population. An example of this effect is the verification of the hypothesis on the significance of the model coefficient of determination (2) for the dependent variable r_4, equation (2.4), which had a clearly unsatisfactory value of $R^2 = 0.1315$ (Table 9): the model variance accounted for only 13.15% of the variance of the dependent variable. With a sample size of 231, the critical significance level of the F-test for R^2 had a p-value $< 10^{-19}$, meaning that the decidedly unsatisfactory value of $R^2 = 0.1315$ was found to be statistically highly significant due to the large sample effect. The criterion of an applicational significant value $R^2 > 0.5$ proved to be valid, while the criterion of statistical significance of the R^2 value proved to be meaningless for assessing the quality of model (2.4). The regression analyses discussed in this publication were conducted for sample sizes greater than 200; for this reason, the negligible critical significance levels of the F-test for $R^2 (p < 10^{-10})$ were not reported. For satisfactory R^2 values above 0.5, statistical significance was automatically achieved. The OLS estimators were fully sufficient for this study, as the consistency between the results of the two transition intensity estimation methods was verified across wide confidence intervals (Table 5, Table 9). It was also confirmed that the MLR and OLS estimators did not differ significantly.

The authors were granted permission to access military vehicle operation records. In peacetime, vehicle usage is, by definition, of extremely low intensity, as these vehicles are designed for wartime operations. To overcome the limitations affecting the reliability of the results – particularly those related to downtime – it would be more beneficial to verify the proposed method using data from a regular vehicle operation process, such as in public transportation or commercial delivery fleets. Only after confirming the methodological effectiveness for regular operation processes could it be applied to the design and optimization of civilian vehicle operation systems. However, such confirmation would require access to civilian databases, such as those used in urban public transport systems.

Discontinuous observations shortened the time intervals of scheduled vehicle operation and reduced the number of observations suitable for reliable estimation of the regression model parameters (2) from 833 to 231. This led to a loss of approximately 72% of the operational process data and nearly doubled the width of the confidence intervals for transition intensities. Eliminating the impact of outages on the results would require repeating the study for a regular vehicle operation process, such as public transport buses. Fluctuations in the frequency of vehicle states were caused by the multi-seasonality of social life, which cannot be avoided in peacetime. These fluctuations contributed to the widening of confidence intervals for transition intensities. Moreover, the low intensity of vehicle operation may have resulted in an underestimation of the frequency of S_5 state observations (vehicle maintenance and repair) compared to the operational design assumptions.

For the five-state operation process studied, there were five state durations and 11 transition intensities (Table 1). According to the factors influencing the Chapman-Kolmogorov predictions (Fig. 2), there was an inevitable and complex interference of each estimated intensity with the other ten when estimating the intensities. Unraveling this interference would require conducting a sensitivity analysis. This task is so difficult that predictions from the Chapman-Kolmogorov system in continuous time are considered unreliable, and the evolution of the facility operation process is instead predicted directly from the distributions of state durations.

ACKNOWLEDGEMENTS

The authors acknowledge the Military University of Technology for financial support (research project UGB 000007-W100-22/2025) in this study.

REFERENCES

[1] F. Zhu *et al.*, "Life cycle optimization framework of charging–swapping integrated energy supply systems for multi-type vehicles", *Appl. Energy*, vol. 351, p. 121759, 2023, doi: 10.1016/j.apenergy.2023.121759.

[2] Z. Cui, X. Wang, Y. Ci, C. Yang, and J. Yao, "Modeling and analysis of car-following models incorporating multiple lead vehicles and acceleration information in heterogeneous traffic flow", *Physica A*, vol. 630, p. 129259, 2023, doi: 10.1016/j.physa.2023.129259.

[3] A. Trivella and F. Corman, "Modeling system dynamics of interacting cruising trains to reduce the impact of power peaks [Formula presented]", *Expert Syst. Appl.*, vol. 230, 2023, doi: 10.1016/j.eswa.2023.120650.

[4] Y. Fu, H. Li, M. Huang, and H. Xiao, "Bi-Objective Modeling and Optimization for Stochastic Two-Stage Open Shop Scheduling Problems in the Sharing Economy", *IEEE Trans. Eng. Manage.*, vol. 70, no. 10, pp. 3395–3409, 2023, doi: 10.1109/TEM.2021.3095954.

[5] D. Rath, S. Kumar, and P.C. Padhi, "On Statistical process control study for front door outer panel of a passenger vehicle through process improvement", *International Journal on Interactive Design and Manufacturing*, vol. 17, no. 3, pp. 1399–1407, 2023, doi: 10.1007/s12008-022-01178-4.

[6] T. Dziubak, "Experimental studies of dust suction irregularity from multi-cyclone dust collector of two-stage air filter", *Energies*, vol. 14, no. 12, p. 3577, 2021, doi: 10.3390/en14123577.

[7] T. Dziubak and L. Bąkała, "Computational and experimental analysis of axial flow cyclone used for intake air filtration in

internal combustion engines", *Energies*, vol. 14, no. 8, p. 2285, 2021, doi: 10.3390/en14082285.

[8] G.V. Loch and A.C.L. da Silva, "A computational study on the number of iterations to solve the transportation problem", *Appl. Math. Sci.*, no. 89–92, pp. 4579–4583, 2014, doi: 10.12988/ams.2014.46435.

[9] Y. Shi *et al.*, "A simulation-optimization system for recycling logistics network of recyclable express packaging", *Comput. Ind. Eng.*, vol. 189, p. 109949, Mar. 2024, doi: 10.1016/j.cie.2024.109949.

[10] H. Liu, S. Li, X. Huang, P. Ding, and Z. Jiang, "Adaptive stochastic configuration network ensemble for structural reliability analysis", *Expert Syst. Appl.*, vol. 237, p. 121633, 2024, doi: 10.1016/j.eswa.2023.121633.

[11] J.Y. Hande and R. Sadiwala, "Optimization of energy consumption and routing in MANET using Artificial Neural Network", *J. Integr. Sci. Technol.*, vol. 12, no. 1, 2024.

[12] Z. Liu, P. Zhao, J. Cao, J. Zhang, and Z. Chen, "A constrained multi-objective evolutionary algorithm with Pareto estimation via neural network", *Expert Syst. Appl.*, vol. 237, p. 121718, 2024, doi: 10.1016/j.eswa.2023.121718.

[13] C. Peng, X. Ying, S. ShanQi, and F. ZiYun, "Application of non-Gaussian feature enhancement extraction in gated recurrent neural network for fault detection in batch production processes", *Expert Syst. Appl.*, vol. 237, p. 121348, 2024, doi: 10.1016/j.eswa.2023.121348.

[14] Y. Wang, L. Guo, M. Wen, and Y. Yang, "Availability analysis for a multi-component system with different k-out-of-n: G warm standby subsystems subject to suspended animation", *Eksploat. Niezawodn.*, vol. 21, no. 2, pp. 289–300, 2019, doi: 10.17531/ein.2019.2.14.

[15] J. Rymarz, A. Niewczas, and A. Krzyżak, "Comparison of operational availability of public city buses by analysis of variance", *Eksploat. Niezawodn.*, vol. 18, no. 3, pp. 373–378, 2016, doi: 10.17531/ein.2016.3.8.

[16] T. Dziubak, T. Wysocki, and S. Dziubak, "Selection of vehicles for fleet of transport company on the basis of observation of their operational reliability", *Eksploat. Niezawodn.*, vol. 23, no. 1, pp. 184–194, 2021, doi: 10.17531/EIN.2021.1.19.

[17] A.C. Cavalea *et al.*, "Maintaining trauma center operational readiness during a pandemic", *American J. Disaster Med.*, vol. 16, no. 1, pp. 25–34, 2021, doi: 10.5055/ajdm.2021.0383.

[18] F. Nasrfard, M. Mohammadi, and M. Rastegar, "Probabilistic optimization of preventive maintenance inspection rates by considering correlations among maintenance costs, duration, and states transition probabilities", *Comput. Ind. Eng.*, vol. 173, p. 108619, Nov. 2022, doi: 10.1016/j.cie.2022.108619.

[19] C.-H. Tang, "A replacement schedule model for airborne service helicopters", *Comput. Ind. Eng.*, vol. 64, no. 4, pp. 1061–1073, 2013, doi: 10.1016/j.cie.2013.02.001.

[20] L. Wang, Q. Yang, and Y. Tian, "Reliability Analysis of 6-Component Star Markov Repairable System with Spatial Dependence", *Math. Probl. Eng.*, vol. 2017, 2017, doi: 10.1155/2017/9728019.

[21] D. Dubový, J. Famfulík, and M. Richtár, "Determination of Operational Reliability of Firefighting Special Vehicles", *Transportation Research Procedia*, 2021, pp. 126–132. doi: 10.1016/j.trpro.2021.06.014.

[22] K. Durczak, P. Rybacki, and A. Sujak, "Application of a Selected Pseudorandom Number Generator for the Reliability of Farm Tractors", *Appl. Sci.*, vol. 12, no. 23, p. 2452, 2022, doi: 10.3390/app122312452.

[23] I.F. Galiev, A.E. Sabitov, and I.F. Galiev, "Analysis of the Reliability and Efficiency of Local Power Supply Systems at Major International Events", in *Lecture Notes in Civil Engineering*, vol. 190, 2022, pp. 269–278, doi: 10.1007/978-3-030-86047-9_28.

[24] J. Kumar, S.A. Bansal, and M. Mehta, "Reliability Analysis – A Critical Review", in *Lecture Notes in Mechanical Engineering*, 2022, pp. 205–217, doi: 10.1007/978-981-16-3135-1_22.

[25] W. Wang, C. Fang, P. Si, Y. Wang, and M. Lin, "Reliability analysis of interval-valued multi-state sliding window system for sequential tasks", *Comput. Ind. Eng.*, vol. 188, p. 109924, Feb. 2024, doi: 10.1016/j.cie.2024.109924.

[26] P. Englezos, "Technology readiness level of gas hydrate technologies", *Can. J. Chem. Eng.*, vol. 101, no. 6, pp. 3034–3043, 2023, doi: 10.1002/cjce.24673.

[27] J. Lewitowicz, M. Zieja, A. Szelmanowski, and A. Pazur, "New measures of operational readiness for multi-states avionics integrated systems with reduced efficiency", in *Lecture Notes in Electrical Engineering*, vol. 548, 2019, pp. 165–180, doi: 10.1007/978-3-030-11187-8_13.

[28] J. Szkutnik-Rogoż, J. Małachowski, and J. Ziołkowski, "An innovative computational algorithm for modelling technical readiness coefficient: A case study in automotive industry", *Comput. Ind. Eng.*, vol. 176, p. 108942, Feb. 2023, doi: 10.1016/j.cie.2022.108942.

[29] T.N. Hoang *et al.*, "Assessment of availability, readiness, and challenges for scaling-up hypertension management services at primary healthcare facilities, Central Highland region, Vietnam, 2020", *BMC Primary Care*, vol. 24, no. 1, 2023, doi: 10.1186/s12875-023-02092-8.

[30] S. Azadiabad, F. Khendek, and M. Toeroe, "Availability and service disruption of network services: From high-level requirements to low-level configuration constraints", *Computer Stand. Interfaces*, vol. 80, p. 103565, 2022, doi: 10.1016/j.csi.2021.103565.

[31] J. Kováč, I. Gregor, J. Melicherčík, and T. Kuvik, "Analysis of the Operating Parameters of Wood Transport Vehicles from the Point of View of Operational Reliability", *Forests*, vol. 14, no. 7, p. 1511, 2023, doi: 10.3390/f14071511.

[32] P. Wróblewski and W. Lewicki, "A method of analyzing the residual values of low-emission vehicles based on a selected expert method taking into account stochastic operational parameters", *Energies*, vol. 14, no. 21, p. 6859, 2021, doi: 10.3390/en14216859.

[33] V. Aulin, A. Hrinkiv, A. Dykha, M. Chernovol, O. Lyashuk, and S. Lysenko, "Substantiation of diagnostic parameters for determining the technical condition of transmission assemblies in trucks", *East. Europ. J. Enterp. Technol.*, vol. 2, no. 1–92, pp. 4–13, 2018, doi: 10.15587/1729-4061.2018.125349.

[34] K. Andrzejczak and J. Selech, "Quantile analysis of the operating costs of the public transport fleet", *Transp. Probl.*, vol. 12, no. 3, pp. 103–111, 2017, doi: 10.20858/tp.2017.12.3.10.

[35] M. Jacyna *et al.*, "Effectiveness of national transport system according to costs of emission of pollutants", in *Proc. Safety and Reliability: Methodology and Applications – Proceedings of the European Safety and Reliability Conference, ESREL 2014*, 2015, pp. 559–567. doi: 10.1201/b17399-79.

[36] S. Vigneaux, "Database Applications for Server-Centered News Production", *138th SMPTE Technical Conference Technical Papers Program*, 1996, pp. 1–15. doi: 10.5594/M00119.

[37] S. Ghosh and S. Roy, "Performance analysis of database models based on fuzzy and vague sets for uncertain query processing", in *Mathematics and Computer Science*, 2024, vol. 2, pp. 363–383. doi: 10.1002/9781119896715.ch25.

[38] M. Izdebski, P. Gołda, and T. Zawisza, "The Use of Simulation Tools to Minimize the Risk of Dangerous Events on the Airport Apron", in *Lecture Notes in Networks and Systems*, vol. 604 LNNS, 2023, pp. 91–107, doi: 10.1007/978-3-031-22359-4_6.

[39] P. Gołda, T. Zawisza, and M. Izdebski, "Evaluation of efficiency and reliability of airport processes using simulation tools", *Eksploat. Niezawodn.*, vol. 23, no. 4, pp. 659–669, 2021, doi: 10.17531/ein.2021.4.8.

[40] J. Ziółkowski, J. Małachowski, M. Oszczypała, J. Szkutnik-Rogoż, and J. Konwerski, "Simulation model for analysis and evaluation of selected measures of the helicopter"s readiness", *Proc. Inst. Mech. Eng., Part G: J. Aerosp. Eng.*, vol. 236, pp. 2751–2762, 2022, doi: 10.1177/09544100211069180.

[41] N. Al Theeb, H.J. Smadi, T.H. Al-Hawari, and M.H. Aljarrah, "Optimization of vehicle routing with inventory allocation problems in Cold Supply Chain Logistics", *Comput. Ind. Eng.*, vol. 142, p. 106341, 2020, doi: 10.1016/j.cie.2020.106341.

[42] J. Ziółkowski, A. Lęgas, E. Szymczyk, J. Małachowski, M. Oszczypała, and J. Szkutnik-Rogoż, "Optimization of the Delivery Time within the Distribution Network, Taking into Account Fuel Consumption and the Level of Carbon Dioxide Emissions into the Atmosphere", *Energies*, vol. 15, no. 14, p. 5198, 2022, doi: 10.3390/en15145198.

[43] H. Cui, J. Qiu, J. Cao, M. Guo, X. Chen, and S. Gorbachev, "Route optimization in township logistics distribution considering customer satisfaction based on adaptive genetic algorithm", *Math. Comput. Simul.*, vol. 204, pp. 28–42, 2023, doi: 10.1016/j.matcom.2022.05.020.

[44] F. Huang, Y. Chen, Y. Chen, and H. Sun, "Stochastic collocation for optimal control problems with stochastic PDE constraints by meshless techniques", *J. Math. Anal. Appl.*, vol. 530, no. 1, p. 127634, 2024, doi: 10.1016/j.jmaa.2023.127634.

[45] N. Isaac and A.K. Saha, "Analysis of refueling behavior of hydrogen fuel vehicles through a stochastic model using Markov Chain Process", *Renew. Sustain. Energy Rev.*, vol. 141, p. 110761, 2021, doi: 10.1016/j.rser.2021.110761.

[46] T. Tang, G. Qin, B. Zhang, J. Cheng, and J. Cao, "Event-based asynchronous state estimation for Markov jump memristive neural networks", *Appl. Math. Comput.*., vol. 473, p. 128653, Jul. 2024, doi: 10.1016/j.amc.2024.128653.

[47] A. Behbahaninia, M. Banifateme, M.H. Azmayesh, S. Naderi, and G. Pignatta, "Markov and monte carlo simulation of waste-to-energy power plants considering variable fuel analysis and failure rates", *J. Energy Res. Technol. Trans. ASME*, vol. 144, no. 6, 2022, doi: 10.1115/1.4051760.

[48] B. Wu and L. Cui, "Reliability evaluation of Markov renewal shock models with multiple failure mechanisms", *Reliab. Eng. Syst. Saf.*, vol. 202, p. 107051, 2020, doi: 10.1016/j.ress.2020.107051.

[49] J. Ziółkowski and A. Borucka, "Markov model in logistic management of enterprise", *J. Konbin*, vol. 38, no. 1, pp. 271–290, 2016, doi: 10.1515/jok-2016-0027.

[50] A. Borucka, "Method of testing the readiness of means of transport with the use of semi-markov processes", *Transport*, vol. 36, no. 1, pp. 75–83, 2021, doi: 10.3846/transport.2021.14370.

[51] E. Kozłowski, A. Borucka, and A. Świderski, "Application of the logistic regression for determining transition probability matrix of operating states in the transport systems", *Eksploat. Niezawodn.*, vol. 22, no. 2, pp. 192–200, 2020, doi: 10.17531/ein.2020.2.2.

[52] J. Ziółkowski, J. Małachowski, M. Oszczypała, J. Szkutnik-Rogoż, and A. Lęgas, "Modelling of the Military Helicopter Operation Process in Terms of Readiness", *Defence Sci. J.*, vol. 71, no. 5, pp. 602–611, 2021, doi: 10.14429/DSJ.71.16422.

[53] H. Vandecasteele and G. Samaey, "Computational efficiency study of a micro-macro Markov chain Monte Carlo method for molecular dynamics", *Appl. Math. Comput.*, vol. 474, p. 128683, Aug. 2024, doi: 10.1016/j.amc.2024.128683.

[54] S. Ghadimi and W.B. Powell, "Stochastic search for a parametric cost function approximation: Energy storage with rolling forecasts", *Eur. J. Oper. Res.*, vol. 312, no. 2, pp. 641–652, 2024, doi: 10.1016/j.ejor.2023.08.003.

[55] C. García-Aroca, M. Asunción Martínez-Mayoral, J. Morales-Socuéllamos, and J.V. Segura-Heras, "An algorithm for automatic selection and combination of forecast models", *Expert Syst. Appl.*, vol. 237, 2024, doi: 10.1016/j.eswa.2023.121636.

[56] J. Ziółkowski, *Stochastyczne modelowanie i optymalizacja gotowości w procesie eksploatacji śmigłowców. Podstawy teoretyczne, metodologia tworzenia, weryfikacja i implementacja modeli*. Warszawa: Wojskowa Akademia Techniczna, 2019 (in Polish).

[57] W.R. Bell, S.H. Holan, and T.S. McElroy, *Economic Time Series: Modeling and Seasonality*. Chapman and Hall/CRC, 2012.

[58] *Skrypt do przedmiotu Ekonometria 1*. M. Rubaszek, Ed., Warszawa: SGH, 2020.

[59] P.A. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation*. New York: John Wiley & Sons, 2017.

[60] Ross, Sheldon M., *Introduction to probability and statistics for engineers and scientists (4th ed.)*. Associated Press, 2009.

[61] *Goodness-of-Fit Techniques*. R.B. D'Agostino, M.A Stephens, Eds., CRC Press, 2020.

[62] P.Ch. Hansen, V. Pereyra, and G.Scherer, *Least Squares Data Fitting with Applications*. Johns Hopkins University Press, 2013.