Evaluation of the maintenance system readiness using the semi-Markov model taking into account hidden factors

Edward Kozłowski\(^a\)\(^*,\), Anna Borucka\(^b\), Piotr Oleszczuk\(^a\), Tomasz Jałowiec\(^c\)

\(^a\) Lublin University of Technology, Poland
\(^b\) Military University of Technology, Poland
\(^c\) War Studies University, Poland

### Highlights
- A competitive model for predicting the readiness of the maintenance system has been developed using the semi-Markov model.
- The method of using the Semi-Markov model in a complex system has been presented.
- The method of estimating the parameters of the semi-Markov model has been presented in a situation where the sojourn time distributions in the given state are not identifiable using one of the classical distributions.
- Diagnostics and evaluation of a transport company in terms of its readiness have been made.

### Abstract
Modelling the time that the system remains in a given state using classical distributions is not always possible. In many cases, empirical distributions are multimodal due to the influence of external, hidden factors and the selection of the best classical distributions may lead to erroneous results. In the article, the method of diagnosis of influence of hidden factors into sojourn time of semi-Markov models was presented. In order to capture hidden factors, the authors proposed to model the distributions of the sojourn time with a mixture of distributions, which is a significant novelty in relation to the studies presented in the literature. Hidden factors directly affect the reliability of technical systems. Detecting the existence of these factors enables more accurate modeling of system readiness. Paying attention to irregularities caused by hidden factors makes it possible to reduce system maintenance costs. Such a system model provides complete information and enables a reliable assessment of the system readiness and maintenance.

### Keywords
maintenance, semi-Markov model, hidden factors, system readiness

### 1. Introduction
Maintaining the proper level of readiness and reliability is crucial for any operating system [33, 44]. It enables the purposeful use of the maintenance potential in the utilization subsystem and periodic reconstruction of this potential in the renewal subsystem in order to maintain the facility's ability to continue operation. This need results from the characteristic features of each technical object immersed in any maintenance system, which include, first of all, limited usability (limited maintenance potential), finite durability and maintenance, material and energy, information and other needs [49].

Therefore, the renewal process is fundamental in any system of operation. It is strongly determined by the adopted maintenance strategy [28]. A strategy focused on preventive maintenance is popular in this regard, used primarily in production systems [37, 72], transport systems [2, 20] or power supply systems [42, 23]. The advantage of this method is its simplicity, as it only requires precise planning of the maintenance schedule [26, 73] without needing specialist knowledge or detailed information about the operation of a specific facility. It is only necessary to know its life cycle (preventive prophylaxis can be based on the passage of a specific time or the performance of a specific job), and the
potential risk of damage [13]. The maintenance plan therefore provides a balance between the risk of failure and the saving of maintenance resources and costs [54]. The disadvantage of this method is primarily unexpected damage, which can cause significant costs, long downtimes and, consequently, delays in the implementation of tasks [69, 45]. Therefore, the answer to these problems is a predictive strategy [11], aimed at optimal use of facilities by eliminating unnecessary periods of downtime resulting from repairs or maintenance [62, 8, 35]. However, it requires detailed information on the current technical condition of the equipment and the implementation of preventive measures [77, 55], as it is based on predicting future conditions of facilities and taking appropriate repair, maintenance and preventive measures in advance.

Therefore, a preventive strategy requires appropriate diagnostic and predictive methods. Popular in this area are classic reliability evaluation methods related to the calculation of basic indicators such as MTBF (mean time between failures) or MTTR (mean time to repair) [31, 50, 22]. More advanced tools are also used, including deep learning methods [9, 38, 10, 53, 77], based on Monte Carlo simulation [46, 21] as well as kriging and first-order reliability [48]. Perti nets [17], Fault tree analysis (FTA) [34], and the load duration distribution method (LDD) [47] are also used. Particularly popular are also time series methods, including e.g. ARIMA class models [19, 27, 24, 32, 65, 67].

However, in time series, the interval between observations is predetermined, therefore modelling the operational states of the object using time series is not a rational operation [7] because they occur at different moments in time. For diagnostics and modelling, the Markov and Semi-Markov models are most often used, where the sojourn time in the given state is a random variable with a specific distribution.

The literature on the application of Markov and Semi-Markov models is extensive. It concerns mainly the evaluation of readiness and reliability of machines and devices in transport engineering or logistics [4, 25, 70]. The authors most often use the theory of one-dimensional Markov and semi-Markov processes in relation to single elements of complex technical objects. An example is the Markov model-based fault diagnosis of wind power converter systems [30], resilience assessment for TLP under mooring failure, the quantitative evaluation of system reliability of flux switching permanent magnet machines (FSPM) [39] or the modular multi-level converter (MMC) [75]. There are definitely fewer works on complex technical objects. One can quote, for example, [36] where a method for assessing the reliability of a wind turbine based on the Hidden-Markov model was proposed, where model of airborne redundant systems operating with faults is developed, or [41] where a method for predicting failure of virtual machines as an object affecting the reliability of cloud platforms based on the AdaBoost-Hidden Markov model was presented. In [3], on the other hand, the reliability of distribution systems was evaluated, taking into account the spatial and temporal distribution of electric vehicles. Studies of complex systems are not popular, however, comprehensive analyses are extremely important, as not only do they allow probing about the reliability of the system as a whole, but also allow to formulate and optimize maintenance policy, which is presented in the publication by Chen and Trivedi [12], Lisnianski and Frankel [43] or Fallahnezhad et al. [18].

To describe a system of operation one must analyze both transition between states in which the system remain (the transition matrix) as well as the time spent by the system in these states (the sojourn time distribution i.e. the time interval between adjacent moments when the stochastic system changes state). For this purpose, we use semi-Markov model in our research.

It is also worth emphasizing that the use of the Markov, Semi-Markov or Hidden-Markov model requires the fulfilment of the assumptions authorizing their use. This is noted, for example, in [4, 76, 64, 15]. These include, first of all, the fulfilment of the Markov property. Only when sequences of random variables satisfy the Markov property can the prediction be accurate [76]. Therefore, it is necessary to test the randomness of the sequence of collected statistical data, as is done, for example, by [71, 61]. The second, important requirement is the assumption regarding the form of distribution of sojourn time in the given states, also emphasized in the literature [40, 53]. The authors use parametric Weibull [68, 15], Poisson [40, 16] and double exponential distributions [60]. Evaluation of the consistency of the distribution of the analysed variable affects the selection of a specific model. For exponential distributions it is a Markov model, and for other
In order to capture hidden factors, the authors proposed to model the distributions of the sojourn time with a mixture of distributions, which is a significant novelty in relation to the studies presented in the literature. The presence of a mixture of decomposition shows that there are additional external factors affecting the analysed state of the object/system, which in this case are taken into account in estimating the level of readiness or evaluating the average sojourn time in this state. Such a system model provides complete information and enables a reliable assessment of the system.

The article consists of four sections. The first, the introduction, contains a description of the problem of modeling the readiness of the maintenance system. Next section is devoted to the presentation of the modelling methods used in the paper. We present the necessary concepts of Markov and semi-Markov processes (basic definitions, identification of semi-Markov processes, estimation of transition probability matrix for Markov process, Markov property test, estimation of sojourn time of system in state) and a method of modelling the sojourn time. Section Results contains the research results of modelling the maintenance system of the police cars performing patrols and interventions in Warsaw, Poland. The last section contains a summary and conclusions.

2. Materials and Methods

2.1. Markov processes

Let \((Ω, F, P)\) be a probabilistic space, \(ℕ\) be the set of natural numbers, \(ℕ₀ = ℕ \cup \{0\}\), \(ℝ\) denotes the set of real numbers and \(S\) be the states space of the analysed system (object, phenomenon). For the transport system we usually assume, that the set \(S\) is finite or countable.

**Definition 1** A sequence of random variables \(\{Xₜ\}_{t∈T}\), \(Xₜ:Ω → S\) for any \(t ∈ T ⊂ ℝ\) is called a stochastic process [29, 57].

If the set of moments \(T\) is finite or countable then the stochastic process is said to be in discrete time (or discrete time stochastic process - DTSP). If \(T\) is uncountable (some subset of set of real numbers), then time is said to be continuous and process is called continuous time stochastic process (CTSP) [29, 57].

We assume that the set of states \(S = \{s₁, s₂, ..., sₖ\}\) (set of possible realisations of the stochastic process \(\{Xₜ\}_{t∈T}\) is finite and \(k ∈ ℕ, 2 ≤ k < ∞\). At any moment \(t ∈ T\) the system can take one of possible realizations and \(Xₜ(ω) = xₜ \in S\). Denote \(P(Xₜ = sᵢ) = pᵢ(𝑡) ≥ 0\), where \(∑ᵢ=1^k pᵢ(𝑡) = 1\), the probability that system is in a state \(sᵢ \in S\), \(1 ≤ i ≤ k\) at the moment \(t ∈ T\).

**Definition 2** Continuous-time stochastic processes \(\{Xₜ\}_{t∈T}\) is called a Markov process \((29, 57)\) if for any \(n ∈ ℕ\), moments \(t₁, t₂, ..., tₙ \in T\) satisfying the condition \(t₁ < t₂ < ... < tₙ\), and states \(xᵢ, x₂, ..., xₙ \in S\), the following property is satisfied:

\[
P(X_{tn} = xₙ|X_{tn₋₁} = xₙ₋₁, X_{tn₋₂} = xₙ₋₂, ..., X_{t₁} = x₁) = P(X_{tn} = xₙ|X_{tn₋₁} = xₙ₋₁) \quad (1)
\]

According to the definition of the Markov process, it follows that the conditional distribution of the random variable \(X_{tn}\) for a given realization sequence \(\{xᵢ\}_{₁≤i≤n₋₁}\) depends only on the last known \(xₙ₋₁\) system state. The property given by the formula (1) is called Markov property or memoryless property, since realization of the stochastic process at the moment \(tₙ\) depends only on the state at the moment \(tₙ₋₁\) but does not depend on previous states (i.e. states at the moments \(tₙ₋₂, ... , t₁\)). For continuous time Markov process it is assumed that the sojourn time \(τₙ = tₙ - tₙ₋₁\) is exponentially distributed.

In the paper we analyze discrete time stochastic process. The process \(\{Xₜ\}_{t∈T}\) at the moment \(t = tₙ\) changes the state and takes the realization \(sᵢ \in S\) and remains in this state until the next transition moment \(tₙ₊₁\), i.e. \(Xₜ = sᵢ\) for \(t ∈ [tₙ, tₙ₊₁)\) and \(Xₜ₊₁ = sᵢ \neq sᵢ\). In other words, at each moment \(tₙ₊₁, n ∈ ℕ\) we observe the jump from the state \(sᵢ\) to different state \(sᵢ\). From above, we take \(Xₜ = Xₜ\) for any \(n ∈ ℕ\) and we accept \(T = ℕ\).

**Definition 3** Discrete-time stochastic processes \(\{Xₜ\}_{t∈N}\) is...
called a Markov chain ([29, 57]), if for any \( n \in \mathbb{N} \) and states \( x_1, x_2, ..., x_n \in S \) the following property is satisfied

\[
P(X_n = x_k | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_1 = x_1) = P(X_{n+1} = x_k | X_n = x_1)
\]

For a heterogeneous Markov chain \( \{X_n\}_{n \in \mathbb{N}} \) the transition probability from state \( s_i \) at the moment \( n \) to the state \( s_j \) at moment \( n + 1 \)

\[
P(X_{n+1} = s_j | X_n = s_i) = p_{ij}(n)
\]

depends on the moment \( n \), \( 1 \leq i, j \leq k \). The conditional distribution of stochastic process depends on the moment in which the process is observed and current state, regardless of previous states. The matrix \( P(n) = [p_{ij}(n)]_{1 \leq i, j \leq k} \) satisfying the condition \( \sum_{j=1}^{k} p_{ij}(n) = 1 \) for \( n \in \mathbb{N} \) and \( p_{ii} = 0 \), \( 1 \leq i \leq k \) is called the (one step) transition probability matrix of Markov chain from \( n \) moment to \( n + 1 \) moment [29, 57].

If the transition probabilities \( p_{ij}(n) \) do not depend on the moment \( n \in \mathbb{N} \) (i.e. \( p_{ij}(n) = p_{ij} \) for \( 1 \leq i, j \leq k \) and any moment \( n \in \mathbb{N} \)) the sequence \( \{X_n\}_{n \in \mathbb{N}} \) is called a homogeneous Markov chain. For homogeneous Markov chain we denote one step transition probability matrix as \( P = [p_{ij}]_{1 \leq i, j \leq k} \), where

\[
\sum_{j=1}^{k} p_{ij} = 1 \quad \text{for} \quad 1 \leq i \leq k \quad \text{and} \quad p_{ii} = 0.
\]

For a homogeneous Markov chain, the transition probability from state \( s_i \) at the moment \( n \) to state \( s_j \) at the moment \( n + m \)

\[
P(X_{n+m} = s_j | X_n = s_i) = [p_{ij}^{(m)}],
\]

where \( [p_{ij}^{(m)}]_{1 \leq i, j \leq k} = P^m \) is the matrix of transition probability in \( m \) steps and \( m \in \mathbb{N} \).

**Definition 4** If stochastic process \( \{X_n\}_{n \in \mathbb{N}} \) is the homogeneous Markov chain and there is a distribution \( \pi = (\pi_1, \pi_2, ..., \pi_k) \), where \( \pi_i \geq 0 \), \( 1 \leq i \leq k \), \( \sum_{i=1}^{k} \pi_i = 1 \) and

\[
\pi P = \pi,
\]

then the distribution \( \pi \) is called the stationary distribution of the homogeneous Markov chain [29, 57].

From (3) - (4) we see that if at certain moment the system has a stationary distribution \( \pi \), then after \( m \in \mathbb{N} \) steps the distribution of states is this same. An important role in the studying of Markov chains is played by its limit properties, especially the boundary distributions \( p_f(n) \) and \( p^*(n) \), \( n \to \infty \).

State \( s_i \) is said to be \( d \) periodic if \( p^*(n) = 0 \) when \( n \) is not divisible by \( d \), and \( d \) is the largest integer with this property.

State \( s_i \) is recurrent if the process starting from this state and returns to this same state in finite time. Aperiodic and recurrent state is called ergodic [29, 57]. The probabilistic properties of the homogeneous Markov chain after a long time (\( m \) transitions, \( m \to \infty \)) are presented in the following theorem.

**Theorem 1** ([29]) Let \( \{X_n\}_{n \in \mathbb{N}} \) be a homogeneous Markov chain with ergodic states and transition probability matrix \( P \in [0,1]^{k \times k} \), then:

a) there is a state probability vector \( \pi = (\pi_1, \pi_2, ..., \pi_k) \) such that \( \pi_i \geq 0 \), \( 1 \leq i \leq k \) and \( \sum_{j=1}^{k} \pi_j = 1 \);

b) \( \pi_j = \lim_{m \to \infty} p_{ij}^{(m)} \) for any \( 1 \leq i, j \leq k \), where

\[
[p_{ij}^{(m)}]_{1 \leq i, j \leq k} = P^m, \quad m \in \mathbb{N}.
\]

c) the stationary distribution \( \pi \) is the solution of equation (4).

The stationary property means that the homogeneous Markov chain being in any state \( s_j \) after a large number of transitions (i.e. \( m \to \infty \)) reaches a stationary distribution \( \pi \) independent of the initial state. To determine the stationary distribution we usually apply the spectral expansion method, see e.g. [29].

### 2.2. Semi-Markov processes

Below we present an extension of discrete time Markov process to its continuous counterpart called semi-Markov process (SMP) [29]. On the one hand we describe the system behavior like a Markov process, on the other we analyze the sojourn time of states of the system.

Let \( \{X_t\}_{t \geq 0} \) be a right-continuous, piecewise constant process, where \( X_t = X_{t_n}(\omega) \) for \( t \in [t_n, t_{n+1}) \) and \( X_{t_n}(\omega) : \Omega \to S \) at transition moments \( t_1, t_2, ..., n \in \mathbb{N} \). The value \( \tau_n = t_n - t_{n-1} \) for \( n \in \mathbb{N} \) denotes \( n \)-th sojourn time and we assume \( \tau_0 = 0 \).

**Definition 5** (Semi-Markov process) A right-continuous, piecewise constant process \( \{X_t\}_{t \geq 0} \) is semi-Markov process, if:

1. the sequence \( \{X_{t_n}\}_{n \in \mathbb{N}_0} \) is homogeneous Markov chain with transition probability matrix \( P = [p_{ij}]_{1 \leq i, j \leq k} \) where \( p_{ij} = P(X_{t_n} = s_j | X_{t_{n-1}} = s_i) \) is the probability of transition from state \( s_i \) to state \( s_j \) at time \( t_n \), \( n \in \mathbb{N} \) and \( \sum_{i=1}^{k} p_{ij} = 1 \);

2. the sojourn (transition) time distribution between states depends on current state and future state (after observing the
jump), thus the distribution matrix of sojourn times

\[ F(t) = \left[ F_{ij}(t) \right]_{1 \leq i,j \leq k} \]

where \( F_{ij}(t) = P(\tau_n \leq t | X_{t_{n-1}} = s_i, X_{t_n} = s_j) \), 1 \( \leq i,j \leq k \) is a distribution function of \( n \)-th sojourn time given that the system obtained the state \( s_i \) at moment \( t_{n-1} \) and shall jump to state \( s_j \) at moment \( t_n = t_{n-1} + \tau_n \).

Semi-Markov process is characterized as embedded Markov chain \( \{X_{tn}\}_{n \in \mathbb{N}} \) with probability transition matrix \( P \) and corresponding to it the sojourn time process \( \{\tau_n\}_{n \in \mathbb{N}} \). Thus the semi-Markov process we present as pair process \( (X_n, t_n)_{n \in \mathbb{N}_0}, \)

\[ X_{t_n} = X_n \text{ for } n \in \mathbb{N}_0 \text{ and } t_n = t_0 + \sum_{i=1}^{n} \tau_i. \]

Below we will estimate distribution of sojourn time that the system at any moment \( t_{n-1} \) take the state \( s_i \) and remains in this state until the next state change. Denote as \( t^i \) the random variable of sojourn time in state \( s_i \). Thus the conditional distribution of sojourn time is the probability that for any \( n \in \mathbb{N} \) the system remains during period \( t_n \) in the state \( s_i \), \( 1 \leq i \leq k \) which has been achieved at time \( t_{n-1} \) as follows

\[ F_i(t) = P(t^i < t) = P(\tau_n < t | X_{t_{n-1}} = s_i) \]

\[ = \frac{1}{P(X_{t_{n-1}} = s_i)} \sum_{j=1}^{k} P(\tau_n < t | X_n = s_j, X_{t_{n-1}} = s_i) \]

\[ = \frac{1}{P(X_{t_{n-1}} = s_i)} \sum_{j=1}^{k} P(\tau_n < t | X_n = s_j, X_{t_{n-1}} = s_i) P(X_n = s_j | X_{t_{n-1}} = s_i) \]

\[ = \sum_{j=1}^{k} F_{ij} P_{ij}. \]

Let \( \{X_i\}_{i \geq 0} \) be a semi-Markov process with embedded homogeneous Markov chain \( \{X_n\}_{n \in \mathbb{N}_0} \) with stationary distribution \( \pi = (\pi_1, \pi_2, ..., \pi_k) \) (corresponding to transition probability matrix \( P \) and satisfying the property (4)) and sequence of conditional distributions of sojourn time \( \{F_i(t)\}_{1 \leq i \leq k}. \) The stationary distribution \( \Pi = (\Pi_1, \Pi_2, ..., \Pi_k), \)

\[ \sum_{i=1}^{k} \Pi_i = 1, \text{ and } \Pi_i \geq 0, 1 \leq i \leq k \text{ of semi Markov process } \{X_i\}_{i \geq 0} \text{ depends on both stationary distribution of embedded Markov chain and conditional distribution of sojourn time } [29]. \]

Let \( Z_i = \{j : X_j(\omega) = s_i\} \), then \( n_i = \# Z_i \) denotes the number of jumps when the embedded Markov chain takes the state \( s_i \) over period \([0, t_{N+1}], N \to \infty. \) From above, the value

\[ \sum_{n \in I} (t_{n+1} - t_n) = \sum_{n \in I} \tau_n \]

is a sojourn time in state \( s_i \) by semi-Markov process over period \([0, t_{N+1}]. \) Then the fraction of time the process spent in state \( s_i \) one can estimate as follows

\[ \eta_j(N) = \frac{\sum_{n \in I} \tau_n}{\sum_{n \in I} \tau_n} \]

From the strong law of large number we have

\[ \frac{1}{N} \sum_{n \in I} \tau_n \xrightarrow{a.s.} E \tau^i \text{ as } N \to \infty. \]

According to definition of fraction time (6) we have

\[ \eta_j(N) \xrightarrow{a.s.} \Pi_i \text{ as } N \to \infty. \]

Finally, substituting (7) - (9) into (6) the stationary distribution of semi-Markov process ([29]) is given by

\[ \Pi_i = \frac{\pi_i E \tau^i}{\sum_{j=1}^{k} \pi_j E \tau^j}. \]

for \( 1 \leq i \leq k. \)

According to definition 5 the identification of semi-Markov model consists of, on the one hand, the estimation of transition probability matrix of embedded homogeneous Markov chain and, on the other hand, the determination of sojourn time distributions between states. After estimation of transition probability matrix we can determine the stationary distribution of Markov chain. From (5) we can estimate the sojourn time distribution in state \( s_i \) and therefore using the formula (10) we determine the stationary distribution of semi-Markov process.

From (5) the sojourn time distribution in state \( s_i \) is defined as linear combination of sojourn time distributions between state \( s_i \) and others. But when we analyse the probability density function (or histogram) of sojourn time in the state \( s_i \) we can observe one or several peaks. These peaks correspond to “latent” factors that can involve sojourn time. The number of peaks may be different from the number of states to which the system can jump (arrive) from the state \( s_i \). In the following, we will present a way to determine the number of possible “latent” factors which involve sojourn time based on kernel density estimation.
2.3. Estimation of transition probability matrix

Let the sequence \( \{x_t\}_{t \in \mathbb{N}} \) denotes realization of the Markov chain and \( x_t \in S = \{s_1, s_2, \ldots, s_n\} \) for \( 0 \leq t \leq n \). To estimate the transition probability matrix for the Markov chain we determine the values \( n_{ij} = |\{t: x_t = s_i, 0 \leq t \leq n\}| \), which is the number of moments when the system remained in the state \( s_i \) for \( 1 \leq i \leq k \), and \( \sum_{i=1}^n n_i = n \). The values \( n_{ij} = |\{t: x_t = s_i, x_{t+1} = s_j, 0 \leq t \leq n - 1\}| \) means the number of transitions from state \( s_i \) to state \( s_j \), \( 1 \leq i, j \leq k \) and \( \sum_{j=1}^n n_{ij} = n_i \). We then determine the estimator of transition probability from state \( s_i \) to state \( s_j \) as \( \hat{p}_{ij} = n_{ij}/n_i \) for \( 1 \leq i, j \leq k \) and the estimated transition probability matrix is equal \( P = [\hat{p}_{ij}]_{1 \leq i,j \leq k} \).

2.4. Markov property test

At the significance level \( \alpha \in (0,1) \) we formulate a null hypothesis:
\[
H_0 : P(X_t = x|X_{t-1} = y, X_{t-2} = z) = P(X_t = x|X_{t-1} = y) \tag{11}
\]
the chain \( \{X_t\}_{t \in \mathbb{N}} \) has a Markov property
and an alternative hypothesis:
\[
H_1 : P(X_t = x|X_{t-1} = y, X_{t-2} = z) \neq P(X_t = x|X_{t-1} = y) \tag{12}
\]
the chain \( \{X_t\}_{t \in \mathbb{N}} \) does not satisfy Markov property, where \( x, y, z \in S \).

To verify the Markov property we apply the \( \chi^2 \) test ([65]). The test statistic is given by
\[
V = \sum_{|i,j| \neq 0, n_{ij} \neq 0} \frac{(n_{ij} - n_{ij}\hat{p}_{ij})^2}{n_{ij}\hat{p}_{ij}}
\]
and has a \( \chi^2 \) distribution with \( k^2(k-1) \) degrees of freedom and denotes differences between \( P(X_t = x|X_{t-1} = y) \) and \( P(X_t = x|X_{t-1} = y, X_{t-2} = z) \) for states \( x, y, z \in S \). Followed by we determine the probability value (p-value, the probability of obtaining test results, [29])
\[
p_{\text{eval}} = \int_{-\infty}^{\infty} \frac{2^{m/2} \Gamma(m/2)}{\pi^{m/2} m^{m/2}} e^{-m |x|^2} dx
\]
where \( m = k^2(k-1) \) and \( \Gamma(\cdot) \) is the gamma function. If \( p_{\text{eval}} > \alpha \) then at significance level \( \alpha \) there are no grounds for rejecting the null hypothesis (we assume that the sequence \( \{X_t\}_{t \in \mathbb{N}} \) satisfies Markov property) otherwise we reject the null hypothesis in favour of the alternative hypothesis (we accept that the sequence does not have Markov property).

2.5. Estimation of sojourn time of system in state

Our method of determination of the probability density function (PDE) estimates of sojourn time in each state was carried out in two steps. In the first step, kernel density estimation (KDE, [6, 51]) for the unknown density function \( f \) was determined. The aforementioned method is one of the nonparametric methods and uses all points of the random sample. It involves “smoothing” the histogram and results in a continuous density function \( \hat{f} \) being an approximation of \( f \). More precisely, let \( X \) be a random variable, and let \( \{x_1, x_2, \ldots, x_n\} \) be its realizations. The kernel density estimator (KDE) of \( f(t) \) is a PDE \( \hat{f}(t) \) having the form
\[
\hat{f}(t) = \frac{1}{n} \sum_{k=1}^n K(x_k, t),
\]
where \( K(x, t), x \in \{x_1, x_2, \ldots, x_n\} \) is a kernel function, being bounded, non negative for all \( x, t \in \mathbb{R} \) and \( \int_{-\infty}^{+\infty} K(x, t)dt = 1 \) for any \( x \in \mathbb{R} \). The kernel \( K \) can be a symmetric or asymmetric function with respect to \( x \). In the case of a symmetric kernel, it can be written in the form
\[
K(x, t) = \frac{1}{h} K\left(\frac{t-x}{h}\right)
\]
where the parameter \( h \), called bandwidth, is responsible for the degree of smoothing. The most commonly chosen symmetric kernels are: the Gaussian, the rectangular, the triangular or Epanechnikov kernel. To determine KDE of sojourn time, the stats::density function of the R language was used, together with Gaussian kernel and ‘rule of thumb’ due to Silverman ([66]) for choosing the bandwidth. Because the KDE is ‘identical’ to the data sample, in the second step, having \( \hat{f}(t) \) i.e. the KDE of \( f(t) \), a parametric PDE from a given family was fitted.

Observing the curve of the kernel density estimator, we see that the curve arises as a mixture of arbitrary densities. In many cases, the hidden factors influence sojourn time. In this case, modelling the sojourn time distribution as a specified distribution does not make sense. The best results we obtain for the sojourn time density being a family of probability distribution functions (PDFs) having the form of convex combination of several densities (hereinafter referred to as a mixture of densities) from a given family i.e.
\[
g(t, c, \alpha, \beta) = \sum_{t=1}^{m} c_j f_d(t, \alpha_j, \beta_j),
\]
where \( c_j \geq 0 \) for \( j = 1,2,\ldots,m \), \( \sum_{j=1}^{m} c_j = 1 \) and \( a = (a_1, a_2, \ldots, a_m) \), \( \beta = (\beta_1, \beta_2, \ldots, \beta_m) \). The number \( m \) in formula (13) is equal to the number of peaks observed in curve of the kernel density estimator \( \hat{f}(t), t > 0 \). We consider the following families of probability density functions ([29]) for \( t > 0, d \in \{W, L, G\} \) forming the mixture (a convex combination (13)):

the random variable \( \tau \) has Weibull distribution, then density is given as follows

\[
f_W(t, a, b) = \frac{a}{b} \left( \frac{t}{b} \right)^{a-1} e^{-\left(\frac{t}{b}\right)^a}
\]

with the shape \( a > 0 \) and scale \( b > 0 \) parameters, and the mean value is equal

\[
E\tau = b \Gamma \left( 1 + \frac{1}{a} \right).
\]

the random variable \( \tau \) has log-normal distribution, then density is given as follows

\[
f_L(t, a, b) = \frac{1}{\sqrt{2\pi b t}} e^{-\frac{(\ln t - a)^2}{2b^2}}
\]

with the mean \( a > 0 \) and the standard deviation \( b > 0 \) and the mean value is equal

\[
E\tau = \exp \left( a + \frac{b^2}{2} \right).
\]

the random variable \( \tau \) has gamma distribution, then density is given by the formula

\[
f_G(t, a, b) = \frac{1}{b^a \Gamma(a)} t^{a-1} e^{-\frac{t}{b}}
\]

with the shape \( a > 0 \), scale \( b > 0 \) parameters, \( \Gamma(\cdot) \) is the gamma function and the mean value is equal

\[
E\tau = ab.
\]

Using the least square method ([6, 27, 24, 65]) we estimate the unknown parameters \( a = (a_1, a_2, \ldots, a_m) \), \( \beta = (\beta_1, \beta_2, \ldots, \beta_m) \) of mixture of the densities (13) i.e. we solve the task

\[
\min_{c,a,\beta} \sum_{i=1}^{n} \left( \hat{f}(x_i) - g(x_i, c, a, \beta) \right)^2,
\]

where \( x_i \) is the sequence of realizations of random variable \( X \).

3. Results

The area that has been selected for research is the maintenance system, in which the readiness of objects to perform tasks is one of the dominant assessment parameters. Such systems include, in particular, state services such as the fire brigade, the police or the armed forces. In this case, police cars were analysed. The subject of the study were cars performing patrol and intervention tasks in the capital city of Poland (Warsaw). The analysis is presented for a selected passenger car. The source database was the documentation of the use of police cars regarding police patrols and registers of technical services and repairs. This made it possible to distinguish the following operational states, in the analyzed sequence:

1. Current Repair (CR) - a car damaged as a result of a breakdown, fault, or involved in a collision event, as a result of which it is unable to perform the task and awaits current / post-accident repair.
2. Repair (Rp) - stoppage in repair.
3. Daily Maintenance 1 (DM1) - daily maintenance is performed to the full extent in the place of permanent or periodic storage of transport equipment, which involves checking the transport equipment, its technical efficiency and technical condition of its assemblies and subassemblies on the day of operation, with particular emphasis on systems that affect driving safety.
4. Daily Maintenance 2 (DM2) - daily maintenance performed to a limited extent, usually after completing the task when patrol activities are completed and the vehicle is parked in the garage.
5. Stand-by (S) – duty stand-by, technically efficient car is parked and ready to carry out the task.
6. Deployment (D) – implementation of the patrol task.
7. Refueling (Rf) – refueling the vehicle.
8. Vehicle Inspection (VI) - technical inspections, in accordance with Polish law, in the case of a new car purchased in a showroom, must be performed within 3 years from the date of first registration. The second inspection should be carried out no later than 2 years after the first one, and the next one every year and the periodic service is performed after the car reaches the inter-service interval or before the winter or summer period.

The table below (tab. 1) presents the transition probability matrix \( P \).
Table 1. The transition probability matrix of Markov chain.

<table>
<thead>
<tr>
<th></th>
<th>CR</th>
<th>DM1</th>
<th>DM2</th>
<th>D</th>
<th>Rf</th>
<th>Rp</th>
<th>S</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0167</td>
<td>0.0000</td>
<td>0.8333</td>
<td>0.0000</td>
<td>0.1667</td>
<td>0.0000</td>
</tr>
<tr>
<td>DM1</td>
<td>0.0011</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.9473</td>
<td>0.0263</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0252</td>
</tr>
<tr>
<td>DM2</td>
<td>0.0095</td>
<td>0.0284</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.9621</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>D</td>
<td>0.0027</td>
<td>0.0108</td>
<td>0.5609</td>
<td>0.2092</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Rf</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0471</td>
<td>0.9333</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0196</td>
</tr>
<tr>
<td>Rp</td>
<td>0.8772</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1228</td>
</tr>
<tr>
<td>S</td>
<td>0.0000</td>
<td>0.9450</td>
<td>0.0000</td>
<td>0.0494</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0056</td>
</tr>
<tr>
<td>VI</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.5833</td>
<td>0.4167</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

The highest values of transition probabilities were achieved for relations with daily service (DM1 and DM2). This is due to the fact that it is an activity partially determined by the instructions in force in most state structures in order to ensure the desired efficiency and readiness of the technical equipment used. It should be carried out each time before commencing a given task - hence the high value of the probability of transition from the state DM1 to D (0.94), and also after its completion, when the car is parked in the garage - hence the high value of the probability of transition from the state DM2 to the state S (0.96) and in the opposite direction from state S to DM1 (0.94). Refuelling the vehicle before carrying out patrols is expressed by a high probability of transition from Rf to D (0.93).

Fig. 1 presents a graphical representation of the transition probability matrix

![Diagram](image)

Figure 1. The interstate transitions graph of Markov chain.

In our case, the value of the statistic (11) is equal to 283.88. The statistic has a $\chi^2$ distribution with 448 degrees of freedom. From (11) the test probability is equal 1. Therefore, at the significance level of 0.05, there are no grounds to reject the null hypothesis, and therefore we assume that the analyzed sequence of states satisfies the Markov property.

For each state, the sojourn time distributions were estimated and fitted using a mixture of densities by solving the task (14). The values of objective function (Sum of Squared Errors) for each fitting are included in Table 9. For some states, the existence of hidden factors affecting the length of stay is clearly visible. The existence of multimodality in empirical distributions (and thus hidden factors) may be caused by the personality features of people operating motor vehicles. Real sojourn times for each state were measured in seconds, but in fitting process we used logarithmic scale for these times.

Starting with the Vehicle inspection state, we notice (fig. 1) that the kernel density estimator has two peaks. For this reason, the combination of 2 densities were fitted. The results of fitting is shown on fig. 2 and related coefficients of convex combination and parameters of fitted distributions are given in the tab. 2 i.e. Weibull 1 and Weibull 2 rows contain coefficients of convex combination and parameters of 2 terms being Weibull distributions (similarly for log-normal and gamma density mixtures).

![Diagram](image)

Figure 2. Sojourn time density for the Vehicle inspection state.

<table>
<thead>
<tr>
<th>Density</th>
<th>j</th>
<th>$c_j$</th>
<th>$\alpha_j$</th>
<th>$\beta_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>1</td>
<td>0.2370</td>
<td>34.2012</td>
<td>9.1430</td>
</tr>
<tr>
<td>Weibull</td>
<td>2</td>
<td>0.7630</td>
<td>22.8309</td>
<td>10.1116</td>
</tr>
<tr>
<td>log-normal</td>
<td>1</td>
<td>0.4611</td>
<td>2.2179</td>
<td>0.0407</td>
</tr>
<tr>
<td>log-normal</td>
<td>2</td>
<td>0.5389</td>
<td>2.3165</td>
<td>0.0341</td>
</tr>
<tr>
<td>gamma</td>
<td>1</td>
<td>0.4795</td>
<td>551.4824</td>
<td>0.0167</td>
</tr>
<tr>
<td>gamma</td>
<td>2</td>
<td>0.5205</td>
<td>901.4588</td>
<td>0.0113</td>
</tr>
</tbody>
</table>

The best result was obtained for the combination of two gamma densities (see: Tab.9). From table 2 we may conclude
that there are two hidden factors directly influencing sojourn time in Vehicle inspection state. One factor has a share of 47.95% and can be identified as random variable with gamma distribution, for which the shape is equal to 551.4824 and the scale - 0.0167, the other has share 52.05% and can be identified as random variable with gamma distribution with shape 901.4588 and scale 0.0113.

The reason may be the inclusion within this state of two types of maintenance, similar in terms of the scope of performed activities, i.e. periodic maintenance and technical inspection. These activities were carried out by various teams (technical inspection by the Police's own workshops and technical inspection by authorized diagnostic stations). The small number of observations related to the technical inspection, which takes place once a year or less, did not justify separating these observations.

In the case of Stand-by state the kernel density estimator has 3 peaks and the combination of 3 densities were fitted. The KDE and fitted parametric densities are shown on fig. 3 and the tab. 3 shows the related coefficients an parameters for terms of fitted distribution mixtures.

The best result for Stand-by state was obtained for the combination of three log-normal densities (see: Tab.9). From table 3 we deduce that there are three hidden factors directly influencing sojourn time. One factor has a share of 14.42% and can be identified as random variable with log-normal distribution, for which the mean is equal to 1.6152 and the standard deviation - 0.2551, the other has share 39.14% and can be described by random variable with log-normal distribution with mean 2.1862 and standard deviation 0.1501, the last one has share 46.44% and log-normal distribution with 2.36352 and 0.0541 parameters.

The KDE for Current repair state has 2 peaks. and the combination of 2 densities were fitted. As before, fig. 4 and the tab. 4 present the fitting results.

Here, the best result was obtain for the combination of two gamma densities (see: tab.9). From table 4 we may conclude that there are two hidden factors directly influencing sojourn time in Current repair state. One factor has a share of 85.31% and can be identified as random variable with gamma distribution, for which the shape is equal to 1404.8564 and the scale - 0.0073, the other has share 14.69% and can be identified as random variable with gamma distribution with shape 216.9024 and scale 0.0554.

This state is characterized by high variability resulting from sudden damage to the vehicle as a result of a breakdown, fault, or participation in a collision event, as a result of which the car

<table>
<thead>
<tr>
<th>Density</th>
<th>$j$</th>
<th>$c_j$</th>
<th>$\alpha_j$</th>
<th>$\beta_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>1</td>
<td>0.1251</td>
<td>5.9973</td>
<td>6.6996</td>
</tr>
<tr>
<td>Weibull</td>
<td>2</td>
<td>0.8749</td>
<td>39.5768</td>
<td>10.3304</td>
</tr>
<tr>
<td>log-normal</td>
<td>1</td>
<td>0.8594</td>
<td>2.3286</td>
<td>0.0268</td>
</tr>
<tr>
<td>log-normal</td>
<td>2</td>
<td>0.1406</td>
<td>2.4877</td>
<td>0.0612</td>
</tr>
<tr>
<td>gamma</td>
<td>1</td>
<td>0.8531</td>
<td>1404.8564</td>
<td>0.0073</td>
</tr>
<tr>
<td>gamma</td>
<td>2</td>
<td>0.1469</td>
<td>216.9024</td>
<td>0.0554</td>
</tr>
</tbody>
</table>

The best result for Stand-by state was obtained for the
is unable to perform the task.

For Daily maintenance 1 state the KDE has 3 peaks and the combination of 3 densities were fitted. fig. 5 and tab. 5 show fitted densities together with KDE and related coefficients and parameters respectively. Here, the best result was obtain for the combination of 3 gamma densities (see: tab.9).

For Daily maintenance 1 the best result was obtain for the combination of three gamma densities (see: tab.9). From table 5 we may conclude that there are three hidden factors directly influencing sojourn time. One factor has a share of 31.49% and can be identified as random variable with gamma distribution, for which the shape is equal to 32.0103 and the scale - 0.1977, the other has share 45.34% and can be identified as random variable with gamma distribution with shape 149.0389 and scale 0.0472, and the last factor has share 23.18% where shape and scale parameters of gamma distribution are equal to 486.5975 and 0.0173 respectively.

For Daily maintenance 2 the best result was obtain for the combination of two gamma densities (see: tab.9). From table 6 we may conclude that there are two hidden factors directly influencing sojourn time. One factor has a share of 41.15% and can be identified as random variable with gamma distribution, for which the shape is equal to 53.6766 and the scale - 0.1149, the other has share 58.85% where shape and scale parameters of gamma distribution are equal to 145.3693 and 0.0497 respectively.

Two distinct review peaks stem from the needs this activity generates. As a rule, it is the same check, however, in the event of deficiencies, correction is necessary. For example, daily maintenance includes checking the cleanliness of the vehicle and, if necessary, washing and cleaning it, checking and, if necessary, topping up the engine oil, or the condition of the coolant, brake fluid, windshield washer fluid, tire pressure, etc. The Daily maintenance 2 state is different: the KDE has 2 peaks. and the combination of 2 densities were fitted. Fig. 6 and the tab. 6 present the fitting results.
tab. 7 present the fitting results. In the case of log-normal and gamma mixtures, the fitted mixture of densities have only 3 terms (instead of 4, one of coefficients of convex combination was close to zero).

![Fitting of density for the Deployment state](image)

**Figure 7.** Sojourn time density for the Deployment state.

**Table 7.** The estimators of mixture of densities the Deployment state.

<table>
<thead>
<tr>
<th>Density</th>
<th>$j$</th>
<th>$c_j$</th>
<th>$\alpha_j$</th>
<th>$\beta_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>1</td>
<td>0.0793</td>
<td>11.7443</td>
<td>6.2013</td>
</tr>
<tr>
<td>Weibull</td>
<td>2</td>
<td>0.2552</td>
<td>11.8686</td>
<td>9.5856</td>
</tr>
<tr>
<td>Weibull</td>
<td>3</td>
<td>0.3026</td>
<td>73.2738</td>
<td>10.1303</td>
</tr>
<tr>
<td>Weibull</td>
<td>4</td>
<td>0.3629</td>
<td>53.1521</td>
<td>10.5063</td>
</tr>
<tr>
<td>log-normal</td>
<td>1</td>
<td>0.3865</td>
<td>2.2657</td>
<td>0.1242</td>
</tr>
<tr>
<td>log-normal</td>
<td>2</td>
<td>0.4646</td>
<td>2.3170</td>
<td>0.0185</td>
</tr>
<tr>
<td>log-normal</td>
<td>3</td>
<td>0.1489</td>
<td>2.3576</td>
<td>0.0105</td>
</tr>
<tr>
<td>gamma</td>
<td>1</td>
<td>0.4263</td>
<td>70.8914</td>
<td>0.1386</td>
</tr>
<tr>
<td>gamma</td>
<td>2</td>
<td>0.4508</td>
<td>3100.8107</td>
<td>0.0033</td>
</tr>
<tr>
<td>gamma</td>
<td>3</td>
<td>0.1229</td>
<td>15071.1378</td>
<td>0.0007</td>
</tr>
</tbody>
</table>

For Deployment state the best fitting was obtain for the combination of four Weibull densities (see: tab.9). From table 7 we may conclude that there are four hidden factors directly influencing sojourn time. One factor has a share of 7.93% and can be identified as random variable with Weibull distribution, for which the shape is equal to 11.7443 and the scale - 6.2013, the second has share 25.52% where shape and scale parameters of Weibull distribution are equal to 11.8686 and 9.5856 respectively, the third has share 30.26% where shape and scale parameters of Weibull distribution are equal to 73.2738 and 10.1303 respectively, the last one has share 36.29% where shape and scale parameters of Weibull distribution are equal to 53.1521 and 10.5063 respectively.

Hidden factors may reflect different types of tasks performed by Police patrols. The time distribution in this state is imposed by the shift work mode of the Police and in accordance with the Act of 6 April 1990 on the Police [1], the shift should be set in a way that allows a police officer to perform his official task within a 40-hour working week, which translates into an 8-hour service. In fact, officers often work in extended shifts, which is caused, among the others, by emergency calls. It should also be remembered that after the end of one shift, another begins, which results in 24-hour police work and continuous operation of vehicles.

In the case of Refueling state, the KDE has 1 peaks and the pure Weibull, log-normal and gamma densities were fitted. The KDE and fitted parametric densities are shown on fig. 8 and the tab. 8 contains the related coefficients and parameters of fitted distributions.

![Fitting of density for the Refueling state](image)

**Figure 8.** Sojourn time density for the Refueling state.

**Table 8.** The estimators of mixture of densities the Refueling state.

<table>
<thead>
<tr>
<th>Density</th>
<th>$j$</th>
<th>$c_j$</th>
<th>$\alpha_j$</th>
<th>$\beta_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>1</td>
<td>1</td>
<td>25.0125</td>
<td>4.9683</td>
</tr>
<tr>
<td>log-normal</td>
<td>1</td>
<td>1</td>
<td>1.5828</td>
<td>0.0465</td>
</tr>
<tr>
<td>gamma</td>
<td>1</td>
<td>1</td>
<td>472.6240</td>
<td>0.0103</td>
</tr>
</tbody>
</table>

For Refueling state the best result was obtain for the Weibull density (see: tab.9). The sojourn time for this state can be identified as random variable with Weibull distribution with the shape parameter 25.0125 and the scale parameter 4.9683.

The refuelling time depends on the capacity of the fuel tank in the car and the efficiency of the dispenser at the petrol station, where the standard refuelling speed is 40 litres per minute. Short/quick refuelling, probably topping up the fuel tank, were more frequent, while longer refuelling in the case of an almost empty tank was less frequent, probably after long patrols or interventions.
The value of the objective function (Sum of Squared Errors) for the task (14) and the sample size for each of the states are given in the table below (tab. 9).

Table 9. The number of observations and SSE for mixtures for the states.

<table>
<thead>
<tr>
<th>State</th>
<th>nr.</th>
<th>n</th>
<th>Weibull mixture</th>
<th>log-normal mixture</th>
<th>gamma mixture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle inspection</td>
<td>12</td>
<td>0.1749</td>
<td>0.0471</td>
<td>0.0315</td>
<td></td>
</tr>
<tr>
<td>Stand-by</td>
<td>891</td>
<td>0.0699</td>
<td>0.0181</td>
<td>0.0202</td>
<td></td>
</tr>
<tr>
<td>Current repair</td>
<td>60</td>
<td>1.5214</td>
<td>0.2783</td>
<td>0.2648</td>
<td></td>
</tr>
<tr>
<td>Daily maintenance 1</td>
<td>873</td>
<td>0.0158</td>
<td>0.0175</td>
<td>0.0158</td>
<td></td>
</tr>
<tr>
<td>Daily maintenance 2</td>
<td>635</td>
<td>0.0767</td>
<td>0.3753</td>
<td>0.0627</td>
<td></td>
</tr>
<tr>
<td>Deployment</td>
<td>1109</td>
<td>0.0443</td>
<td>0.3591</td>
<td>0.4486</td>
<td></td>
</tr>
<tr>
<td>Refueling</td>
<td>255</td>
<td>4.2083</td>
<td>5.0563</td>
<td>11.3582</td>
<td></td>
</tr>
</tbody>
</table>

For the Repair state we have two point distribution. Once the state was reached, the system moved on to the other state after 57600 or 230400 sec (in logarithm scale 10.961 and 12.348).

The sojourn time for this state is presented on fig. 9.

Figure 9. Sojourn time density for the Repair state.

The specificity of the repair status results from the repair possibilities and the availability of parts. The car was repaired immediately after finding the fault or waited for repair for a maximum of two days. Hence the two-point distribution of this state.

Based on results presented in tab. 10 we estimate the stationary distribution \( \{ \pi_i \}_{1 \leq i \leq 8} \) of Markov chain which satisfies the property (4). From (13) we determine densities for each state and choose such type of density for which the objective function of task (14) has the smallest value. Followed by we assess the expected sojourn time \( E \tau_i \) for each state, \( 1 \leq i \leq 8 \). From (10) we estimate the stationary distribution of semi-Markov process \( \{ \Pi_i \}_{1 \leq i \leq 8} \). We present the results in tab. 11.

Table 10. Stationary distribution of Markov chain, expected sojourn times in states, stationary distribution of semi-Markov process.

<table>
<thead>
<tr>
<th>nr.</th>
<th>State</th>
<th>Distribution</th>
<th>( \pi_i )</th>
<th>( E \tau_i )</th>
<th>( \Pi_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vehicle inspection</td>
<td>gamma mixture</td>
<td>0.0031</td>
<td>9.7005</td>
<td>0.0036</td>
</tr>
<tr>
<td>2</td>
<td>Stand-by</td>
<td>log-normal mixture</td>
<td>0.2291</td>
<td>9.2155</td>
<td>0.2563</td>
</tr>
<tr>
<td>3</td>
<td>Current repair</td>
<td>gamma mixture</td>
<td>0.0154</td>
<td>10.5228</td>
<td>0.0197</td>
</tr>
<tr>
<td>4</td>
<td>Daily maintenance 1</td>
<td>gamma mixture</td>
<td>0.2242</td>
<td>7.1364</td>
<td>0.1942</td>
</tr>
<tr>
<td>5</td>
<td>Daily maintenance 2</td>
<td>gamma mixture</td>
<td>0.1631</td>
<td>6.7938</td>
<td>0.1345</td>
</tr>
<tr>
<td>6</td>
<td>Deployment</td>
<td>Weibull mixture</td>
<td>0.2849</td>
<td>9.6285</td>
<td>0.3329</td>
</tr>
<tr>
<td>7</td>
<td>Refueling</td>
<td>Weibull mixture</td>
<td>0.0655</td>
<td>4.8612</td>
<td>0.0386</td>
</tr>
<tr>
<td>8</td>
<td>Repair</td>
<td>two point</td>
<td>0.0147</td>
<td>11.2773</td>
<td>0.0201</td>
</tr>
</tbody>
</table>

The \( \Pi_i \) values obtained make it possible to evaluate the limit probability of the tested system in certain states. The highest value (33%) was obtained for the state of Deployment, which means that vehicles are used in 1/3 of the time. Another high indication was for the Stand-by state (25%), meaning a stoppage and waiting for a task. Both of these states are those in which vehicles are ready to perform tasks or are already performing them. The next in line indications concerned the mandatory daily service of vehicles and amounted to 19% and 13%, respectively. Although these are conditions related to the renewal of vehicles, they are rather related to checking the technical condition and inspection before departure, so it can also be considered that vehicles in these conditions are ready to perform the task. The probabilities of being in the other states classified as unready were relatively low and amounted to approximately 4% Refuelling, 2% Repair and Current repair, and only 0.4% Vehicle inspection, respectively. This allows to conclude that the tested system was in almost 92% of the state of readiness, of which, above all, it performed its tasks, which proves the good management of the owned infrastructure. The low rate of unreadiness, amounting to 8%, is satisfactory and allows to conclude about the high technical culture and properly implemented renewal treatments.

4. Conclusion

In the article a transport system identification was done using a semi-Markov model. The transition matrix between states was identified (the Markov property was verified). In addition, a stationary distribution was determined for the Markov chain. Identification of sojourn times in states has been done. The kernel density estimations of sojourn times in states clearly showed that it is not possible to approximate them using...
classical theoretical distributions. The impossibility of the mentioned approximation is due to latent factors. Therefore, the kernel density estimations of sojourn times in states were approximated using a mixture of certain classical distributions: Weibull, log-normal and gamma. As the results showed, for different states the kernel density estimation had to be approximated by mixtures of different distributions. For one of the states, a two-point distribution was identified. The expected sojourn times were then determined for each of the states. The fitted parametric distributions in the form of mixtures allowed to determine the stationary distribution of the semi-Markov process.

On the one hand the identification of stationary distribution of the semi-Markov process and on the other the estimation of distributions of the sojourn times in these states, enable to assess the readiness of the tested system more precisely. System readiness is defined as the total probability when the vehicle is in a state of immediate performance of the task (Deployment), but also in readiness to perform the task (Stand-by) and during vehicle being serviced (Daily Maintenance 1 and Daily Maintenance 2) directly required by the internal regulations of the Police. On the other hand, in the Refuelling, Repair, Current repair and Vehicle Inspection states, the vehicle is taken out of service, so these are non-ready states.

An important achievement of the study presented is the ability to detect hidden factors using a mixture of distributions that are not possible to be captured when analysing the sojourn time using classical distributions. Hidden factors affect the reliability of technical systems, increase the uncertainty of maintenance and have a direct impact on the cost of maintaining the system. Detection of these factors enables more accurate system readiness modeling and optimization of system maintenance.

Additionally the method of identification of sojourn time in states enables to simulate transportation systems because empirical distributions of sojourn times in states have been accurately approximated by mixtures of classical ones.

The results obtained clearly show that the distribution of the sojourn time of objects in individual states is influenced by hidden factors. On this basis, it can be concluded that they correspond to hidden states, therefore, further research will be conducted towards modelling the maintenance system using the Hidden Markov Model, which allows to take into account the sojourn time in the hidden states.

References

32. Kumaresan K, Ganeshkumar P. Software reliability prediction model with realistic assumption using time series (S) ARIMA model. Journal

Eksploatacja i Niezawodność – Maintenance and Reliability Vol. 25, No. 4, 2023


52. Pawlik P, Kania K., Przysucha B., Fault diagnosis of machines operating in variable conditions using artificial neural network not requiring


