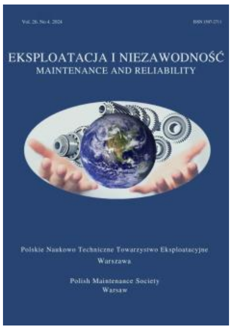

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 26 (2024), Issue 4

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

Article citation info:
Kozłowski E, Borucka A, Oleszczuk P, Leszczyński N, Evaluation of readiness of the technical system using the semi-Markov model with selected sojourn time distributions, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2024: 26(4) http://doi.org/10.17531/ein/191545

Evaluation of readiness of the technical system using the semi-Markov model with selected sojourn time distributions

Indexed by: Web of Science Group

Edward Kozłowski[a,*], Anna Borucka[b], Piotr Oleszczuk[a], Norbert Leszczyński[c]

[a] Lublin University of Technology, Poland
[b] Military University of Technology, Poland
[c] University of Life Sciences in Lublin, Poland

Highlights

- A method for detailed analysis of the exploitation process using Markov process theory has been presented.
- Identification of the sojourn time distribution using selected random variable distributions has been made.
- The proposed approach allows an assessment of the probability of occurrence of the individual states and its duration.
- The model that provides significant support for forest management processes has been made.

Abstract

Mechanization of forestry work is crucial in forest management, and the specific nature of the tasks performed requires reliable machines with a high level of technical readiness. Therefore, models describing the exploitation process of multi-operational machines used to obtain wood raw materials, and especially assessing the level of their technical readiness, are extremely important. For multi-tasking objects performing random activities in given time intervals, calculating readiness measures is a complex issue. Forecasting subsequent operational states and their durations allows (if the Markov property is met) to predict the behavior of technical objects and schedule work. In addition to identifying the sequence of states, it is also important to identify the sojourn time in these states. This article presents a method for identifying the semi-Markov process and then assessing the technical readiness of a Harvester machine. This allowed for conclusions regarding the timely completion of assigned tasks and also made it possible to adjust the activities carried out to the requirements of forest management.

Keywords
technical readiness, reliability, empirical distribution fitting, semi-Markov model, forest management, Harvester logging machines

1. Introduction

Forestry and the timber industry play a crucial role in the economies of many countries, and effective and proper forest exploitation is essential to ensure biodiversity and maintain the continuity of ecosystem functions. Therefore, it is important to maintain a balance between forest cultivation and forestry operations. This also means implementing modern technologies, including the use of the best methods for timber harvesting. Currently, in times of increased mechanization and digitization, there is a dynamic growth in the use of harvesters [2]. Working with such forestry machines is not only 6-8 times more efficient compared to the alternative use of a chainsaw [15], but also increases work safety and ergonomic comfort, as well as positively influences the attitudes of machine operators towards modern work techniques [6]. All of this contributes to conducting forest management in a more sustainable manner. To this end, rigorous periodic inspections of machines working in the forest are also conducted [7]. In addition, daily checks of fluid levels and assessments for fluid leaks, inspections of all

(*) Corresponding author.
E-mail addresses: E. Kozłowski (ORCID: 0000-0002-7147-4903) e.kozlovski@pollub.pl, A. Borucka (ORCID: 0000-0002-7892-9640) anna.borucka@wat.edu.pl, P. Oleszczuk (ORCID: 0000-0002-6515-257X) p.oleszczuk@pollub.pl, N. Leszczyński (ORCID: 0000-0002-5198-8341) norbert.leszczynski@up.lublin.pl,

screw connections, tightening or replacing them, and examining the structure for cracks are required. For these vehicles, daily maintenance is also performed, which involves removing needles, leaves, snow, ice, and other debris that may pose a fire hazard, restrict visibility, or hinder the assessment of technical condition. However, despite such a rigorous system of monitoring the technical condition of machines, they can still break down. This is primarily due to the work cycle of the harvester, which is characterized by a large number of short-term intensive technological operations and is also exposed to adverse environmental impacts. On one hand, this poses a threat to the natural environment, and on the other, it can cause undesirable work interruptions, reducing the profitability of the timber harvesting process. Therefore, methods for assessing technical readiness, especially for such machines, are desirable for both environmental and economic reasons, particularly in terms of predicting malfunctions, identifying wear, or detecting operational irregularities [17]. They also support machine lifespan management, enabling effective equipment lifecycle management and planning for replacement by providing essential information for decision-making regarding potential investments in new machinery or the modernization of existing ones [16, 18]. Additionally, they contribute to optimizing work efficiency and improving safety. Hence, literature contains studies regarding the operation of harvesters. Most of them focus on evaluating machine performance. For instance, in [3], algorithms based on a decision tree, gradient boosting machine, linear regression, k-nearest neighbors, support vector machine, and artificial neural network were utilized to study machine efficiency. Similarly, in [5], based on datasets from forest inventory and the machine fieldbus, the article examined the impact of the forest environment and harvester operation method on machine efficiency using unsupervised machine learning methods (clustering model). For example, in Sweden, the potential of using measurement data collected from cut trees to provide accurate forest estimates at the stand level was presented, thereby supporting decisions at various levels of the forestry industry chain [4]. Similar research is conducted by Liski and others [8], utilizing gradient boosted machine (GBM), support vector machine (SVM), and ordinary least square (OLS) regression to predict the productivity of cut-to-length (CTL) harvesting. In article [5], a linear regression model was proposed to predict machine fuel consumption based on operator input data and forest inventory data. Most studies concerning the monitoring and assessment of mechanized wood harvesting operations using harvesters utilize analytical tools such as machine learning. A comprehensive literature review in this field conducted by Maktoubian et al. [9] demonstrates a significant advantage in the utilization of artificial intelligence in predictive maintenance issues and their continuous development [20]. However, as the authors claim, the quantity and quality of available data limit their use, and the results are not as reliable as expected [3, 8]. The reason for this limitation is the decentralization and lack of data management in the forest environment, which makes it impossible to acquire the required datasets [8]. Due to the obstacles related to the digitization of the forestry sector, arising from both technical and socio-economic factors associated with the development of digital technologies in such areas [10], it is necessary to seek alternative assessment methods. The above became the genesis of this article, in which a semi-Markov model was used to study the technical readiness of heavy forestry machinery such as Harvesters. In this article, it was assumed that the exploitation process of the studied technical object involves the deliberate utilization of its operational potential during task execution and the periodic restoration of this potential to maintain the object's ability to continue operating. Within this framework, five operational states describing the machine's activity and one state assigned to renewal processes were distinguished. Based on empirical data, the semi-Markov process was identified, and the technical readiness was assessed. The duration of the system's stay in states is influenced by many factors, sometimes difficult to precisely identify [18, 22].

In the literature, the duration is most often modeled as a random variable with an exponential distribution. However, classical distributions do not always accurately reflect the properties of the random variable describing the time spent in a state. Therefore, this article proposes the use of alternative theoretical distributions, demonstrating their greater effectiveness in identifying the sojourn time in a state. Such an approach is not popular in the literature primarily due to the difficulty in estimating the parameters of such distributions. Especially when studying real exploitation processes, this is a significant challenge [18, 27, 28].

Additionally, these distributions may vary depending on the sequence of states, an issue often overlooked in the literature. In the study of technical readiness, this is particularly relevant for the state of malfunction, as it significantly affects the outcome. Therefore, the article proposes a method that identifies the distribution of downtime with repair considering hidden factors that influence the duration, namely the probabilities of transitioning to the next state. This makes the proposed method innovative.

The scientific contribution of this article is as follows:

1. Proposing a method for detailed analysis of the exploitation process using Markov process theory.
2. Identifying the distribution of sojourn time in each state using alternative distributions of the random variable.
3. Identifying the distribution of downtime with repair considering the sequence of future states.
4. Justifying the particular importance of the repair state in forest work planning and scheduling.
5. Proposing an approach to assess the probability of occurrence of individual states, especially the repair state, and its duration.
6. Presenting a model that provides significant support for forestry management processes.

The structure of the article is as follows. After introducing and identifying the research gap, the theoretical foundations, limitations, and requirements of the semi-Markov process are presented. Next, a characterization of possible theoretical distributions from the family of exponential distributions is provided, as they are deemed best for assessing equipment operating time.

In the subsequent part of the study, parameters of the semi-Markov model are estimated, determining the transition probability matrix and identifying the distribution of sojourn time in each state. Finally, a characterization of the studied technical object is provided, describing the identified states using selected distributions, allowing for the estimation of the system's technical renewal time.

2. Materials and methods

2.1. Subject of the study

A harvester is a multi-operational and high-performance forestry vehicle with an articulated structure, which is intended primarily for felling trees and wood manipulation. It is used for delimbing (cutting off side branches) of felled trees and cutting their trunks into ready-made assortments. Sometimes it is also used to debark tree trunks. At the same time, it is used to arrange assortment packages along the skidding route, which are later collected by forwarders. Harvesters can achieve capacities ranging from several to even 30 m3/h depending on the vehicle model, availability of trees, and logging system, and are able to cut trees with a diameter of up to 20 or even 102 cm (depending on the size of the machine and the head used). The study used data from the on-board computer of the Ponsse Ergo 8W harvester with an H7 thinning and felling head, with an automatic cutting unit, allowing the cutting of trees with a diameter of up to 72 cm. The tested machine operated in the area of Roztocze Wschodnie in the Tomaszów Forest District, in a mature stand, in a fresh mixed forest with a volume of 372 m3/ha.

Figure 1. Ponsse Ergo 8W Harvester with an H7 thinning and felling head.

The collected empirical data constituting the basis for the research were in the form of PDF files (Portable Document Format). They contained chronological data on subsequent stages (states) of the harvester's work on each working day, including the start time of a given stage, duration, distance covered, fuel consumption, and data on harvested wood. Figure 1 shows a photo of the harvester being the source of empirical data, while Table 1 contains its technical parameters

Table 1. Technical data of Ponsse Ergo 8W Harvester.

Parameter	Unit	Value
Operating Weight	kg	21500
Base Carrier Length	mm	8130
Carrier Width	mm	2630
Ground Clearance	mm	600
Transport Height	mm	3800
Engine Output	kW	205
Engine Torque	Nm	1100
Tractive Force	KN	195
Fuel Tank	l	380
Crane Pump Capacity	cm^3	145
Head Pump Capacity	cm^3	190
Hydraulic Oil Tank	l	290
Wheel Equipment	----	720x45x26,5

2.2. Semi-Markov process

The machines in use are in various states related to their current use or technical condition. In general, the set of these states is finite and will be denoted by $S = \{x_1, x_2, \cdots, x_k\}$. The most commonly used mathematical model of such a machine maintenance process is a stochastic process with values in the set S i.e. a sequence of random variables $\{X_t\}_{t\in\mathbb{R}}$, $X_t: \Omega \to S$ for any $t \in \mathbb{R}$ (continuous-time stochastic process) or $t \in \mathbb{N}$ (discrete-time stochastic process), where $\mathbb{R}$ and $\mathbb{N}$ are the sets of real numbers and natural numbers including zero respectively.

The Markov chains with finite set of states are widely used to model the behavior of technical systems.

Definition 1. A discrete-time stochastic processes $\{X_t\}_{t\in\mathbb{N}}$ satisfying the memoryless property:

$$P\big(X_n = x_{i_n} | X_{n-1} = x_{i_{n-1}}, X_{n-2} = x_{i_{n-2}}, \dots, X_0 = x_{i_0}\big) = P\big(X_n = x_{i_n} | X_{n-1} = x_{i_{n-1}}\big).$$

for any $n \in \mathbb{N}$ and states $x_{i_0}, x_{i_1}, \dots, x_{i_n} \in S$ is called a Markov chain.

An important role in the technical system modelling is played by homogeneous Markov chains for which the transition probability $P\big(X_{n+1} = x_j | X_n = x_i\big) = p_{ij}(n)$ from state $x_i \in S$ to state $x_j \in S$ does not depend on moment n, i.e. $p_{ij}(n) = p_{ij}$ for any $n \in \mathbb{N}$.

Let $P = [p_{ij}]_{i,j=1,2,\dots,k}$ be the transition probability matrix for a homogeneous Markov chain and $P\big(X_{n+m} = x_j | X_n = x_i\big) = p_{ij}^{(m)}$ denotes the transition probability from state x_i to state x_j in m steps. If the initial distribution of X_0 random variable (i.e. probability vector $p(0) = \big(p_1(0), p_2(0), \dots, p_k(0)\big)$, $p_i(0) = P(X_0 = x_i)$ for $x_i \in S$, where $0 \le p_i(0) \le 1$ for $i = 1,2,\dots,k$ and $\sum_{i=1}^{k} p_i(0) = 1$) is known, then the probability vector $p(n) = \big(p_1(n), p_2(n), \dots, p_k(n)\big)$ being the distribution of the random variable X_n (namely the vector of probabilities of being in particular states at the moment n) is determined by the formula:

$$p(n) = p(0)P^n,$$

where

$$\big[p_{ij}^{(m)}\big]_{1\le i,j\le k} = P^m.$$

Definition 2 Let $\{X_t\}_{t\in\mathbb{N}}$ be a homogeneous Markov chain. A distribution $\pi = (\pi_1, \pi_2, \dots, \pi_k)$, where $0 \le \pi_i \le 1$ for $i = 1,2,\dots,k$ and $\sum_{i=1}^{k} \pi_i = 1$ satisfying the equation

$$\pi P = \pi, \tag{1}$$

is called the stationary distribution of the homogeneous Markov chain.

A stationary distribution plays a key role in the study of limit properties of probabilities $p_{ij}^{(m)}$, when $m \to \infty$, which is important in reliability analysis and is the content of the theorem below. First recall that a state s_i is called d periodic if d is the largest integer such that $p_{ii}^{(kd)} > 0$ and $p_{ii}^{(n)} = 0$ for any n not divisible by d. A chain is called irreducible if from each state it is possible to reach every other state after a finite number of steps. A state s_i is called is recurrent if the process starting s_i returns to it in finite time. A state s_i is said to be ergodic if it is aperiodic and recurrent.

Theorem 1 Let $\{X_n\}_{n\in\mathbb{N}}$ be a homogeneous Markov chain with ergodic states and P be its transition probability matrix. There exists a stationary distribution $\pi = (\pi_1, \pi_2, \dots, \pi_k)$, such that

$$\pi_j = \lim_{m\to\infty} p_{ij}^{(m)},$$

for any $1 \le i,j \le k$.

The theorem states that a homogeneous Markov chain with ergodic states, reaches a stationary distribution after a sufficiently long period of time [30].

The semi-Markov process, which we will use in our research, is a continuous time stochastic process and is based on the concept of a Markov chain [19, 21]. This is a generalization of the continuous time Markov process, where the sojourn time distribution depends on the current state and need not be exponential. On the one hand, the behavior of the system is described by a Markov chain, while on the other hand, we analyze the sojourn time in states [18].

Definition 3 (Semi-Markov process) The right-continuous and piecewise constant stochastic process $\{X_t\}_{t\geq 0}$ is semi-Markov process, if:

1. there exist random time moments $t_0 < t_1 < t_2 \ldots$ ($t_n \in \mathbb{R}$, $n \in \mathbb{N}$) such that the sequence $\{X_{t_n}\}_{n\in\mathbb{N}}$ is a homogeneous Markov chain with transition probability matrix $P = [p_{ij}]_{1\leq i,j\leq k}$, where $p_{ij} = P(X_{t_{n+1}} = x_j | X_{t_n} = x_i)$ for $i, j \in \mathbb{N}$;
2. the distribution of the random variable $\tau_n = t_{n+1} - t_n$, $n = 1,2, \ldots , \ldots$, namely the n-th sojourn time given that the system obtained the state x_i at moment t_n and shall jump to state x_j at moment $t_{n+1} = t_n + \tau_n$, depends only on the current state and the future state (after observing the jump) i.e. the probability $P(\tau_n \leq t | X_{t_n} = x_i, X_{t_{n+1}} = x_j)$ depends only on x_i and x_j.

Therefore, a semi-Markov process is a pair process $\{(X_n, t_n)\}_{n\in\mathbb{N}}$, where $X_n = X_{t_n}$ for $n \in \mathbb{N}$ and $t_n = t_0 + \sum_{j=1}^{n} \tau_j$, characterized by both its embedded Markov chain $\{X_{t_n}\}_{n\in\mathbb{N}}$ and corresponding the sojourn time stochastic process $\{\tau_n\}_{n\in\mathbb{N}}$. Hence, the identification of semi-Markov model consists of the estimation of transition probability matrix of $\{(X_n, t_n)\}_{n\in\mathbb{N}}$ and the determination of sojourn time distributions for each state. A semi-Markov process in which the sojourn times for all states are exponentially distributed is called a continuous-time Markov chain [30].

Let τ^i be a random variable denoting the sojourn time in state s_i. The sojourn time distribution in state s_i, $1 \leq i \leq k$ is estimated as follows:

$$F_i(t) = \sum_{j=1}^{k} P\left(\tau_n < t | X_{n+1} = s_j, X_n = s_i\right) \times$$

$$P\left(X_{n+1} = s_j | X_n = s_i\right) = \sum_{j=1}^{k} F_{ij}\,(t) P_{ij}. \qquad (2)$$

The stationary distribution of semi-Markov process $\{X_t\}_{t\geq 0}$ is calculated as

$$\Pi_i = \frac{\pi_i E\tau^i}{\sum_{j=1}^{k} \pi_j\, E\tau^j}, 1 \leq i \leq k, \qquad (3)$$

where the embedded homogeneous Markov chain $\{X_n\}_{n\in\mathbb{N}}$ have a stationary distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_k)$, but $E\tau^i$ denotes expected sojourn time in state s_i, $1 \leq i \leq k$.

2.3. Markov property test

Let $\{x_t\}_{0\leq t\leq n}$ be a realization of the Markov chain, where $x_t \in S$ for $0 \leq t \leq n$ and let $n_i = \#\{t: x_t = s_i, 0 \leq t \leq n\}$ for $i = 1,2, \ldots, k$ k i.e. the number indicating how many times the system remained in the state s_i, obviously $\sum_{i=1}^{k} n_i = n$. The value $n_{ij} = \#\{t: x_t = s_i, x_{t+1} = s_j, 0 \leq t \leq n-1\}$ for $1 \leq i, j \leq k$ denotes the number of transitions from state s_i to state s_j , thus $\sum_{j=1}^{k} n_{ij} = n_i$. Let $\hat{p}_{ij}$ be the estimator of probability p_{ij} (the transition probability form state s_i to state s_j) and $\hat{p}_{ij} = n_{ij}/n_i$ for $1 \leq i, j \leq k$. The estimated transition probability matrix of the embedded homogeneous Markov chain has the form $P = [\hat{p}_{ij}]_{1\leq i,j\leq k}$.

To verify the Markov property, at the significance level $\alpha \in (0,1)$ we formulate a null hypothesis:

$H_0 : P(X_t = x | X_{t-1} = y, X_{t-2} = z) = P(X_t = x | X_{t-1} = y)$ (the chain $\{X_t\}_{t\in\mathbb{N}}$ has a Markov property)

and an alternative hypothesis:

$H_1 : P(X_t = x | X_{t-1} = y, X_{t-2} = z) \neq P(X_t = x | X_{t-1} = y)$ (the chain $\{X_t\}_{t\in\mathbb{N}}$ does not satisfy Markov property),

where $x, y, z \in S$.

The test statistic

$$V = \sum_{i,j,l=1, n_{ij}\neq 0, n_{jl}\neq 0}^{k} \frac{\left(n_{ijl} - n_{ij}\hat{p}_{jl}\right)^2}{n_{ij}\hat{p}_{jl}} \qquad (4)$$

has χ^2 distribution with $(k-1)^2 k$ degrees of freedom. The probability value (the p-value)is given by

$$p_{val} = \int_{V}^{\infty} \frac{x^{m/2-1} e^{-x/2}}{2^{m/2}\Gamma(m/2)} dx \qquad (5)$$

where $m = (k-1)^2 k$ and $\Gamma(\cdot)$ is the gamma function. If $p_{val} \leq \alpha$ then at significance level α we reject the null hypothesis in favor of the alternative hypothesis (we accept that the sequence does not have Markov property) otherwise we have no grounds to reject the null hypothesis and assume that the sequence $\{x_t\}_{0 \leq t \leq n}$ satisfies the Markov property.

2.4. Family of exponential distribution

The family of exponential distributions is an important class of probability distributions used in statistics and probability theory, particularly in reliability analysis [23]. The following probability distributions were used to identify sojourn time in states:

- gamma distribution;
- log-normal distribution;
- Weibull distribution;
- inverse Weibull distribution;
- log-Weibull distribution;
- compound Weibull distribution;
- Keis and Phani's distribution;
- generalized Weibull distribution;
- gamma Weibull distribution.

2.4.1. Gamma distribution

The gamma distribution is widely used in probability theory, statistics, and reliability analysis. It is also defined by two parameters. The gamma distribution with shape $b > 0$ and scale $1/a$ (rate $a > 0$) [26] has density

$$f_G(x) = \frac{a^b}{\Gamma(b)} x^{b-1} e^{-ax}, x > 0, \tag{6}$$

and the cumulative distribution function is given by

$$F_G(x) = \frac{\gamma(b, ax)}{\Gamma(\alpha)}, x > 0. \tag{7}$$

where the gamma function is defined by $\Gamma(s) = \int_0^{+\infty} t^{s-1} e^{-t} dt$ and the incomplete gamma function is defined by $\gamma(s, x) = \int_0^{x} t^{s-1} e^{-t} dt$.

2.4.2. Log-normal distribution

The log-normal distribution is a probability distribution of a random variable whose natural logarithm has a normal distribution. Thus if random variable Y has a normal distribution, then the random variable $X = exp\,(Y)$ has a log-normal distribution. It is used to model behavior of features whose realizations take only positive values. The random variable X has a log-normal distribution with parameters a and b, if the random variable Y has a normal distribution with mean a and standard deviation b. The log-normal distribution with $a \in \mathbb{R}$ and $b > 0$ has the density

$$f_{LN}(x, a, b) = \frac{1}{\sqrt{2\pi} b x} e^{-\frac{(\ln x - a)^2}{2b^2}}, x > 0 \tag{8}$$

and the cumulative distribution function is given by

$$F_{LN}(x, a, b) = F\left(\frac{\ln x - a}{b}\right), x > 0 \tag{9}$$

where F denotes the normal distribution $N(0,1)$.

2.4.3. Weibull distribution

The Weibull distribution with shape $\beta > 0$ and scale $1/\sqrt[\beta]{\alpha}$ ($\alpha > 0$) [24, 25] has density given by

$$f_W(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^\beta}, x > 0, \tag{10}$$

and the cumulative distribution function is given by

$$F_W(x) = 1 - e^{-\alpha x^\beta}, x > 0. \tag{11}$$

2.4.4. Inverse Weibull distribution

The inverse Weibull distribution is most commonly used for modeling extreme values, particularly in the context of reliability and risk analysis. If the random variable Y has a Weibull distribution, then the random variable $X = 1/Y$ has an inverse Weibull distribution. Usually, this distribution is used to describe the time to failure in situations where the risk of failure decreases with time [29]. The inverse Weibull distribution with parameters $a > 0$ and $b > 0$ has density given by

$$f_{IW}(x) = abx^{-b-1} e^{-ax^{-b}}, x > 0, \tag{12}$$

and the cumulative distribution function is given by

$$F_{IW}(x) = e^{-ax^{-b}}, x > 0, \tag{13}$$

2.4.5. Log-Weibull distribution

The log-Weibull distribution [29] is a continuous probability distribution that combines features of the Weibull distribution and the logarithmic transformation. Let Y be the random variable which has Weibull distribution with $\alpha > 0$ and $\beta > 0$ parameters, then the random variable $\frac{X-a}{b} = log\left(\alpha Y^\beta\right)$ has log-Weibull distribution with $a \in \mathbb{R}$ and $b > 0$ and density function is given by

$$f_{LW}(x) = \frac{1}{b} e^{(x-a)/b} e^{-e^{(x-a)/b}}, x \in \mathbb{R}, \tag{14}$$

whereas the cumulative distribution function is given by

$$F_{LW}(x) = 1 - e^{-e^{\frac{(x-a)}{b}}}, x \in \mathbb{R} \tag{15}$$

2.4.6. Compound Weibull distribution

Let $Y_1, Y_2, \dots, Y_N$ be a sequence of independent identically Weibull distributed random variables, N be geometric random variable with parameter p and N is independent of $Y_1, Y_2, \dots, Y_N$. The random variable $X = \min\{Y_1, Y_2, \dots, Y_N\}$ has compound (geometric) Weibull distribution. The density function of compound Weibull distribution [29] is given by

$$f_{WG}(x) = \frac{ab^a(1-p)x^{a-1}e^{-(bx)^a}}{(1-pe^{-(bx)^a})^2}, x > 0, \tag{16}$$

and the cumulative distribution function is given by

$$F_{WG}(x) = \frac{1-e^{-(bx)^a}}{1-pe^{-(bx)^a}}, x > 0, \tag{17}$$

where $a > 0$, $b > 0$ and $0 < p < 1$.

2.4.7. Keis and Phani's (modified Weibull) distribution

The Phani-Keis distribution is a specialized probability distribution that is also used in reliability analysis. Keis introduced a modification of the Weibull distribution by adding the constraint that the realization of a random variable takes values from the interval (a, b), where $0 < a < b$. Keis distribution is defined by shape β and scale α parameters. Phani [31] proposed an extension with an additional shape γ parameter. The density function of modified Weibull distribution is given by

$$f_{PK}(x) = \frac{\alpha(x-a)^{\beta-1}\left((b\beta - a\gamma) + (\gamma - \beta)x\right)}{(b-x)^{\gamma+1}} e^{-\alpha\frac{(x-a)^\beta}{(b-x)^\gamma}}, \tag{18}$$

where $x > 0$, $0 < a < b < \infty$ and $\alpha, \beta, \gamma > 0$. The cumulative distribution function is given by

$$F_{PK}(x) = 1 - e^{-\alpha\frac{(x-a)^\beta}{(b-x)^\gamma}} \tag{19}$$

The special case of modified Weibull distribution $\gamma = \beta > 0$ was presented by Keis [30].

2.4.8. Generalized Weibull distribution

Mudholkar and Kollia [32] proposed an extension of the standard Weibull distribution. The generalized Weibull distribution is used most often to model the lifetime of technical objects, in reliability analysis when ageing or wear processes are relevant. The generalized Weibull distribution is described by three parameters: the shape parameter $b > 0$ determines shape of the survival curve, the scale parameter $a > 0$ determines the time scale and the shift parameter c. Depending on shift parameter c the realization of random variable takes values from the set

$$D = \begin{cases} (0, \infty), & \text{dla } c < 0, \\ (0, (ac)^{-1/b}), & \text{dla } c > 0. \end{cases}$$

The density function of generalized Weibull distribution is given by

$$f_{GW}(x) = abx^{b-1}(1 - acx^b)^{\frac{1}{c}-1}, x \in D, \tag{20}$$

for $a, b > 0$ and $c \neq 0$, the cumulative distribution function is given by

$$F_{GW}(x) = 1 - (1 - acx^b)^{\frac{1}{c}}, x \in D. \tag{21}$$

2.4.9. Gamma Weibull distribution

The gamma-Weibull distribution [33] is a specific combination of the gamma distribution and the Weibull distribution. This makes it possible to fit more flexibly to different empirical data by combining the characteristics of both distributions. The gamma-Weibull distribution is defined by three parameters. For $a > 0$, $b > 0$ and $k > 0$ the density function is given by

$$f_{GW}(x) = \frac{ba^k}{\Gamma(k)} x^{kb-1} e^{-ax^b}, x > 0, \tag{22}$$

and the cumulative distribution function is given by

$$F_{GW}(x) = \frac{\gamma(k, ax^b)}{\Gamma(k)}, x > 0, \tag{23}$$

where $\gamma(a, x)$ denotes the incomplete gamma function. When $k = 1$ we have Weibull distribution, whereas with $b = 1$ we have gamma distribution.

3. Identification of the technical system

3.1. Characteristics of operational states

From chronological readings regarding the harvester's work in the period from March 3, 2021, to June 24, 2022 (a total of 113 working days), a set of data was created regarding subsequent stages (states) of the harvester's work on each working day, taking into account the start time of a given stage and its duration. The first stage of building a Markov model is to isolate the possible operating states of the tested machine. Six states have been specified, which are presented, along with their characteristics, in Table 1. At any time, the machine can be in one of them.

Table 2. Operational states of the tested forest machine.

State name	Characteristics
Break	interruptions due to various factors
Driving	drives to the working surface and along operational routes, lasting longer than 1 minute
Off-road travel	exits from the operational trail (i.e. driving around the operational field), to hard-to-reach trees
Other work	application status update, data transfer, communication, head calibration and others
Processing	drives between working positions (shorter than 1 minute), sequential cutting down of subsequent trees from a given working position and producing various ready-made wood assortments from them (firewood, logs, sawmill wood, etc.)
Repair	waiting time for repair and removal of the fault, transport to the service center and removal of the fault

3.2. Identification of the semi-Markov process

A correct description of the exploitation system still requires defining the relations between its elements, therefore it was necessary to determine possible transitions of the object between individual states, which was presented in the form of a matrix (Table 2.) and using a graph (Figure 1), which also presents the values of the probability of transitions between individual states.

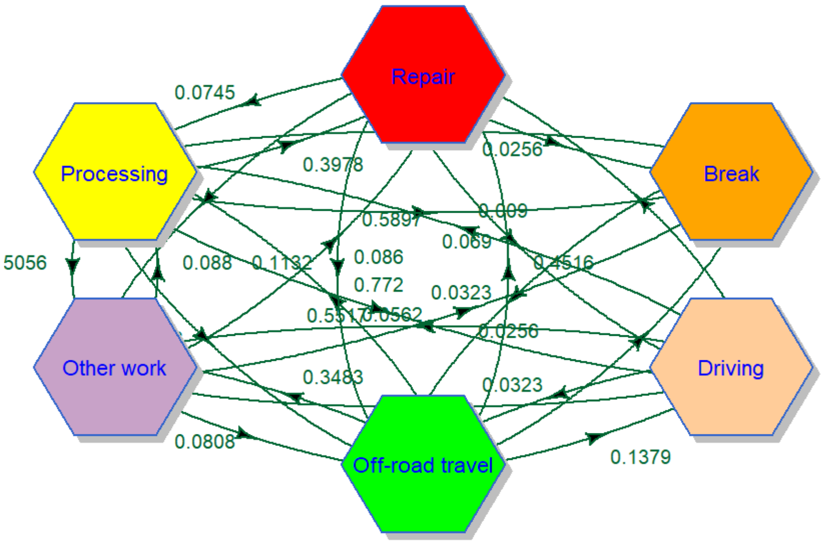

Figure 1. Transition graph between states.

Table 3. Transition probability matrix for a Markov chain.

	Break	Driving	Off-road travel	Other work	Processing	Repair
Break	0.0000	0.0000	0.3590	0.0256	0.5897	0.0256
Driving	0.0000	0.0000	0.1379	0.2414	0.5517	0.0690
Off-road travel	0.0323	0.0323	0.0000	0.0808	0.7413	0.1132
Other work	0.0000	0.0562	0.3483	0.0000	0.5056	0.0899
Processing	0.0564	0.0090	0.7720	0.0880	0.0000	0.0745
Repair	0.0000	0.0645	0.4516	0.0860	0.3978	0.0000

The value $V = 156.95$ of the test statistic was estimated using formula (4). The distribution of the test statistic is χ^2 with 150 degrees of freedom. From formula (5), the test probability equals $p_{val} = 0.3322$. Therefore, at the significance level of 0.05, there are no grounds to reject the working hypothesis, we assume that the sequence of states satisfies the Markov property.

Next, the stationary distribution for the Markov chain, which describes the long-term behavior of the Markov chain when it reaches an equilibrium state, was calculated. This allows to assess the limit behavior of the tested exploitation system. The expected sojourn times in each state were also determined, and then the limit distribution of the semi-Markov process was estimated using formula (3). Stationary distribution and expected sojourn times in the states are presented in Table 4.

Table 4. Stationary distribution for the Markov chain, expected sojourn times in a state, and the stationary semi-Markov distribution.

	π_i	$E\tau^i$	Π_i
Break	0.0345833	28.282051	0.0322402
Driving	0.0257821	47.827586	0.0406458
Off-road travel	0.3843541	8.969977	0.1136429
Other work	0.0798523	32.888889	0.0865677
Processing	0.3927980	47.426637	0.6140604
Repair	0.0826303	41.430107	0.1128430

The obtained results show that almost 39.28% falls in the Processing state, i.e. from the Markov chain properties we can conclude that the tree-cutting machine was in this state with a probability of 39.28%. However, taking into account the average sojourn times and analyzing the stationary distribution of the semi-Markov process, we see that the tested system tends to stay in the Processing state, which is over 61% of the total exploitation of the Harvester machine.

The second-best result concerns off-road travel and amounts to over 11.36%. This is due to the need to move the machine to the place of work. The remaining values are negligible and range from 3% for the Break state, through 4% for the Driving state, and almost 8% for the other work state. Being in the Repair state took over 11% of the total operating time.

4. Repair time analysis

From the point of view of the correctness of the operation process and prevention of potential disruptions, the repair time is the most important, therefore it will be subject to detailed analysis. Predicting repair time is an important element of creating a machine operation schedule. Nevertheless, repair time schedules τ depend on the subsequent states in which the object resides, therefore they are estimated as follows:

$$\begin{aligned} F_{\text{Repair}}(t) &= \sum_{j \in S \setminus \{\text{Repair}\}} P\left(\tau_n < t | X_{n+1} = j, X_n = \text{Repair}\right) P(X_{n+1} = j | X_n = \text{Repair}) \\ &= \sum_{j \in S \setminus \{\text{Repair}\}} F_{\text{Repair},j}\left(t\right) P_{\text{Repair},j}. \end{aligned} \tag{24}$$

Therefore, in order to obtain full information regarding the repair time of a technical object in the operation system, Repair time distributions should be determined, taking into account the transition to the future state. From Table 3 we see that from the Repair state the system moves to the states: Driving, Off-road travel, Other work, Processing states. Depending on the future state, the distribution of sojourn time in the Repair state was identified. The maximum likelihood method was used to identify the parameters of the distributions.

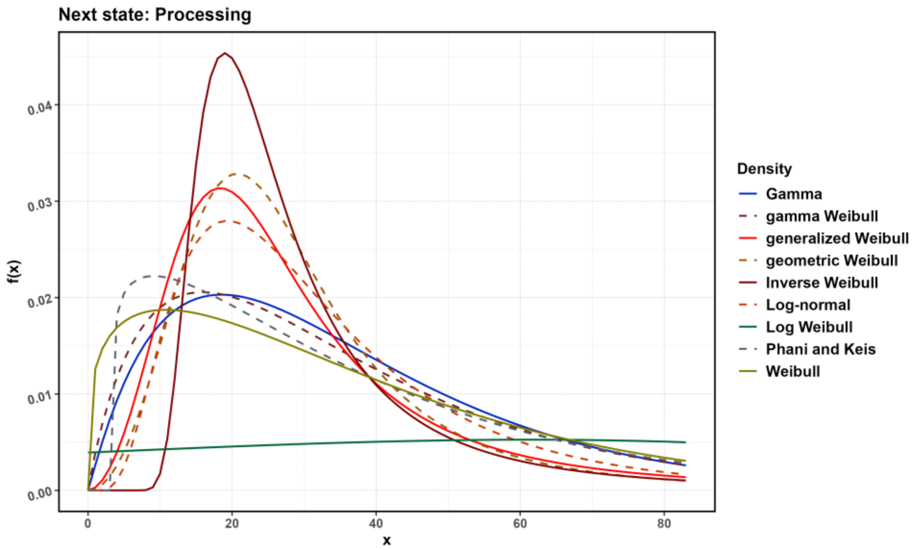

Figure 2. Fitting plots of the sojourn time distribution in the Repair state during the transition to Processing state.

Table 5 shows the parameter values of the distributions of the sojourn time in the Repair state and the value of the logarithm of the likelihood function in the case of transition to the Processing state. Figure 2 shows the density functions for each type of distribution.

Table 5. Parameters of the sojourn time in Repair during the transition to Processing state and the values of the logarithm of the likelihood function.

Type	Parameters	LogLik
Weibull	a = 0.0102 ; b = 1.2458	-168.142
Gamma	a = 0.0561 ; b = 2.0395	-202.155
Log-normal	a = 3.3363 ; b = 0.6121	-157.781
Inverse Weibull	a = 2144.8819 ; b = 2.4924	-150.494
Log Weibull	a = 61.4028 ; b = 69.9605	-207.422
geometric Weibull	a = 3.0464 ; b = 0.004 ; p = 0.999	-156.168
Phani and Keis	a = 3.3718 ; b = 285.6 ; α = 0.0172 ; β = 1.1433 ; γ = 0.000001	-171.101
generalized Weibull	a = 1e-04 ; b = 2.8918 ; c = -1.3086	-155.878
gamma Weibull	a = 0.0881 ; b = 0.9003 ; k = 2.1728	-201.755

The highest value of the likelihood function was obtained for the inverse Weibull distribution. The sojourn time in the Repair state and the transition to the Processing state should be modeled using a random variable with inverse Weibull distribution with parameters a = 2144.8819 ; b = 2.4924.

Table 6 presents the values of parameters of the sojourn time distributions in the Repair state and the value of the logarithm of the likelihood function in the case of transition to the Other work state. Figure 3 shows the density functions for each type of distribution.

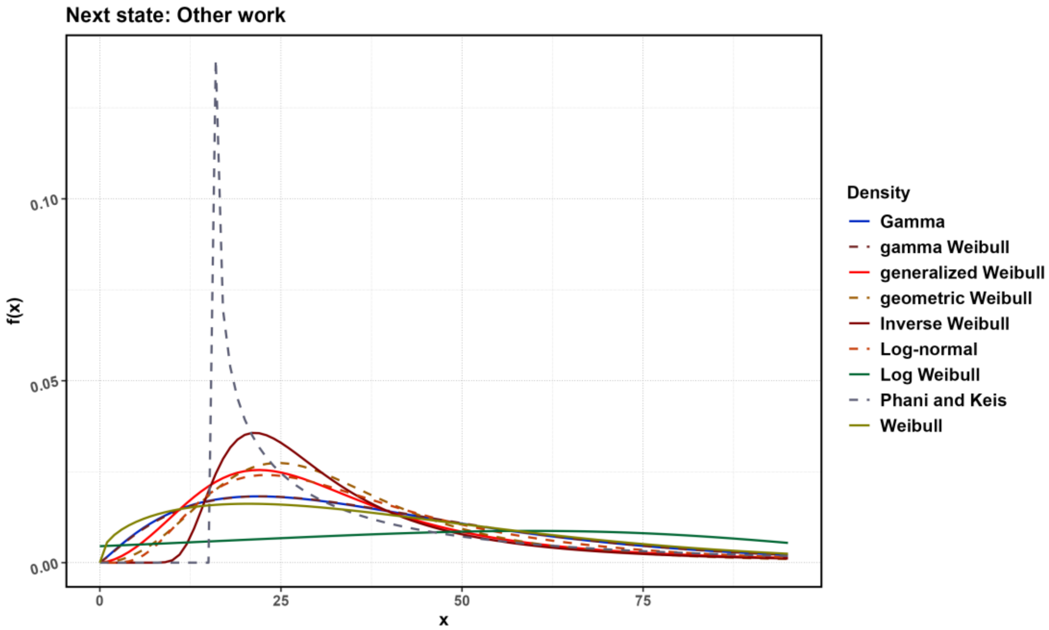

Figure 3. Fitting plots of sojourn time distribution in the Repair state during the transition to the Other work state.

Table 6. Parameters of the sojourn time in Repair during the transition to Other work state and the values of the logarithm of the likelihood function.

Type	Parameters	LogLik
Weibull	a = 0.0039 ; b = 1.4532	-36.773
Gamma	a = 0.0525 ; b = 2.1458	-44.275
Log-normal	a = 3.496 ; b = 0.6019	-35.258
Inverse Weibull	a = 1367.9745 ; b = 2.2383	-34.355

Type	Parameters	LogLik
Log Weibull	a = 59.6117 ; b = 42.1587	-41.487
geometric Weibull	a = 3.0182 ; b = 0.0058 ; p = 0.9945	-35.138
Phani and Keis	a = 15.84 ; b = 147.6 ; α = 0.1384 ; β = 0.7022 ; γ = 0.0329	-35.397
generalized Weibull	a = 1e-04 ; b = 2.7202 ; c = -1.1517	-35.208
gamma Weibull	a = 0.0496 ; b = 1.0126 ; k = 2.1327	-44.274

Table 7 presents the values of the parameters of sojourn time distribution in the Repair state and the value of the logarithm of the likelihood function in the case of transition to the Off-road travel state. Figure 4 shows the density functions for each type of distribution.

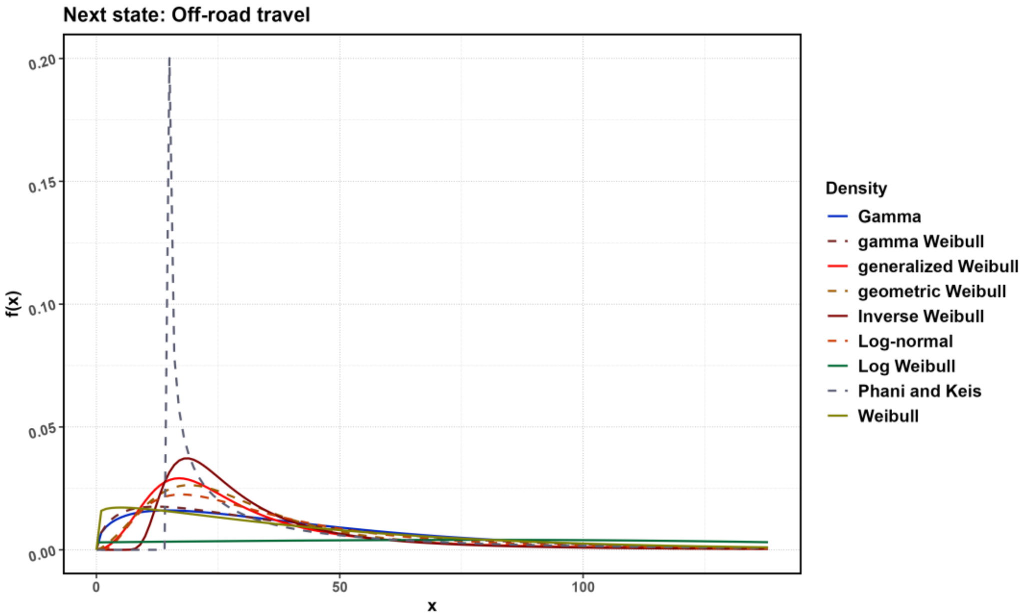

Figure 4. Fitting plots of sojourn time distribution in the Repair state during the transition to Off-road travel state.

The highest value of the likelihood function was obtained for the Phani and Keis distribution. The sojourn time in the Repair state and the transition to the Off-road travel state should be modeled using Phani and Keis distribution with parameters a = 14.85 ; b = 373.2 ; α = 0.1648 ; β = 0.6051 ; γ = 0.00001.

Table 7. Parameters of sojourn time in Repair during the transition to Off-road travel state and the values of the logarithm of the likelihood function.

Type	Parameters	LogLik
Weibull	a = 0.0148 ; b = 1.0931	-201.875
Gamma	a = 0.0325 ; b = 1.4747	-242.811
Log-normal	a = 3.4372 ; b = 0.76	-192.433
Inverse Weibull	a = 603.2954 ; b = 2.0566	-184.079
Log Weibull	a = 78.7882 ; b = 90.6824	-246.793
geometric Weibull	a = 2.4207 ; b = 0.0021 ; p = 0.999	-191.684
Phani and Keis	a = 14.85 ; b = 373.2 ; α = 0.1648 ; β = 0.6051 ; γ = 0.00001	-178.26
generalized Weibull	a = 1e-04 ; b = 2.9284 ; c = -1.8947	-188.496
gamma Weibull	a = 0.1225 ; b = 0.7709 ; k = 2.1474	-240.838

Table 8 presents the values of the parameters of sojourn time distributions in the Repair state and the value of the logarithm of the likelihood function in the case of transition to the Driving state. Figure 5 shows the density functions for each type of distribution.

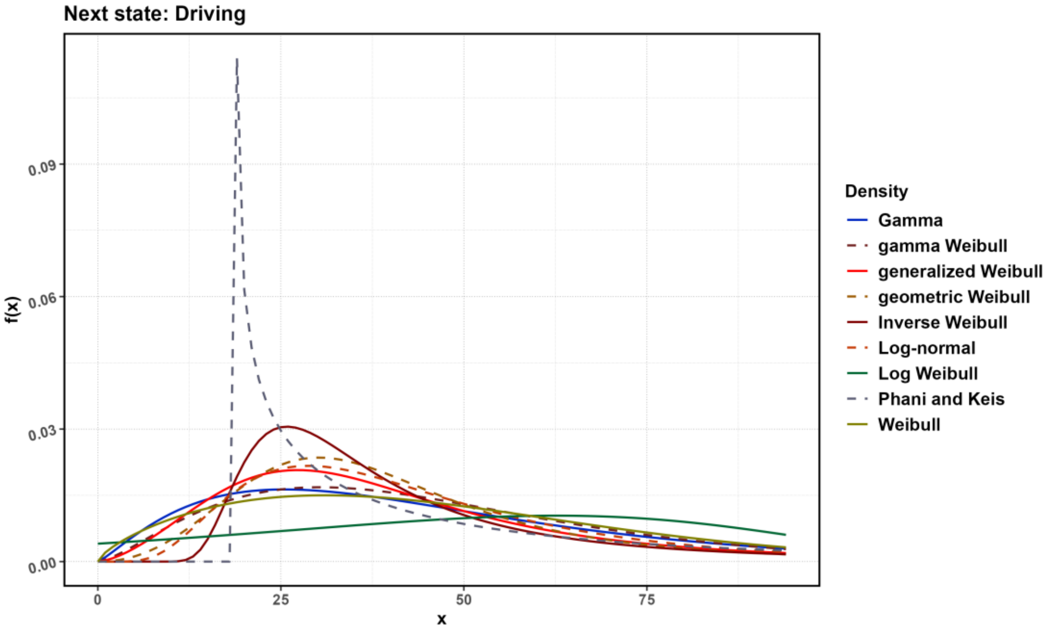

Figure 5. Fitting plots of sojourn time distribution in the Repair state during the transition to Driving state.

Table 8. Parameters of sojourn time in Repair during the transition to Driving state and the values of the logarithm of the likelihood function.

Type	Parameters	LogLik
Weibull	a = 0.001192 ; b = 1.702891	-27.757
Gamma	a = 0.0487 ; b = 2.2415	-33.574
Log-normal	a = 3.6617 ; b = 0.5499	-26.896
Inverse Weibull	a = 2603.1611 ; b = 2.3081	-26.474
Log Weibull	a = 62.6339 ; b = 35.37828	-30.218
geometric Weibull	a = 3.1325 ; b = 0.00768 ; p = 0.9803	-26.928
Phani and Keis	a = 18.81 ; b = 130.8 ; $\alpha = 0.142839$; $\beta = 0.709003$; $\gamma = 0.070675$	-27.744
generalized Weibull	a = 1e-04 ; b = 2.5459 ; c = -0.8647	-27.142
gamma Weibull	a = 0.0237 ; b = 1.168 ; k = 2.1098	-33.467

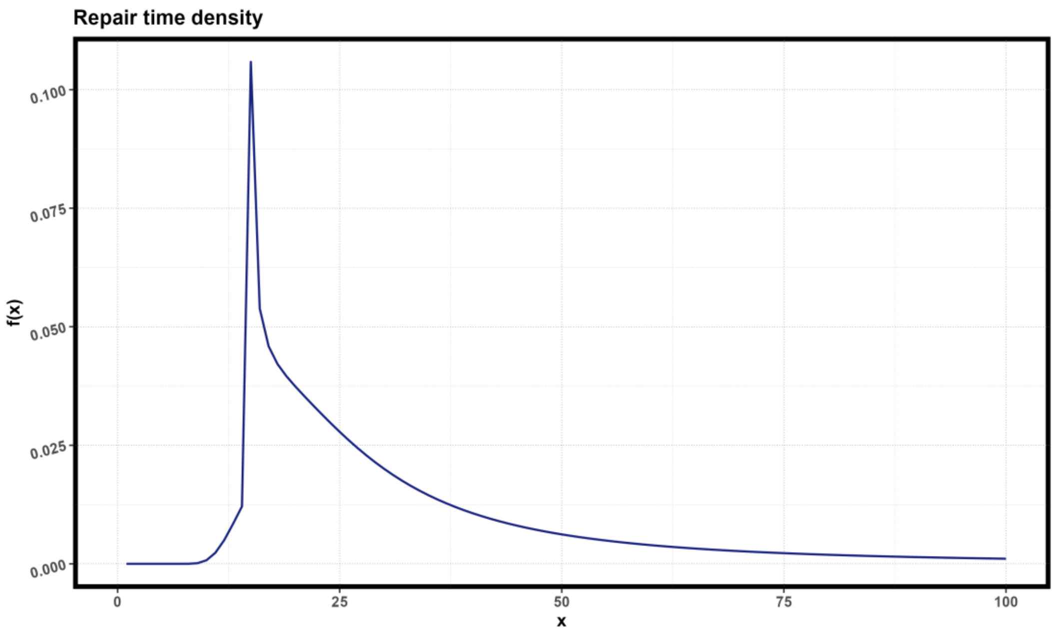

Figure 6. Sojourn time distribution in the Repair state.

The highest value of the likelihood function was obtained for the inverse Weibull distribution. The sojourn time in the Repair state and the transition to the Off-road travel state should be modeled by random variable having inverse Weibull distribution with parameters a = 2603.161 ; b = 2.3081.

In the next stage of the study, using the results obtained above (Table 4-7), the repair time distribution was estimated using formula (24). The density function is shown in Figure 6.

The expected repair time is estimated using the formula

$$E\tau^{Repair} = \int_0^\infty t \frac{\partial}{\partial t} F_{\text{Repair}}(t)$$

and $E\tau^{Repair} = 34.53$ min.

This value is lower than the classically estimated value presented in Table 3. This is due to the fact that in Table 3 estimated the expected values of the stay time in states without taking into account the change in the sequence of states.

Moreover, the identification of the repair time taking into account the transition to next states resulted in a change in the stationary distribution of the semi-Markov process, which is presented in Table 8.

Table 8. Stationary distribution of the semi-Markov process after taking into account the change in the expected sojourn time in the Repair state

	Break	Driving	Off-road travel	Other work	Processing	Repair
Π_i	0.03286	0.04142	0.11582	0.08823	0.62582	0.09586

After correcting the time spent in the Repair state, we see that the analyzed system stays in the Processing state for more than 62% of the total utilization, while in the Repair state, it spends less than 10% of the total utilization. Taking into account the probability of transition between states allowed for a more accurate identification of the time distribution of the sojourn time in the repair state. For each transition from the repair state to other states, an appropriate sojourn time distribution was determined, and among those presented, the distributions with the highest value of the likelihood function were selected. A more accurate estimate of the repair time distribution enables better identify the stationary distribution of the system modeled by semi-Markov process.

5. Conclusion

The application of semi-Markov process theory allowed for a detailed analysis of the exploitation process, in this case, concerning a forestry machine such as a Harvester. Special attention was given to the repair state, as it is crucial in planning and scheduling forest work. Through analysis using Markov processes, the repair time distribution was presented as the scalar product of the repair time distribution and the transition probability between the repair state and the other states. In the literature on the subject the use of known random distributions or mixture of distributions to describe operational events is often used and is not new [18,27,28]. The approach presented in the paper consists on solution the problem of estimation of readiness of technical system. The originality of presented method is the identification of the machine's behavior performing many tasks in variable operating conditions, which enable the modeling work plans. Therefore, less common distributions of random variables in the literature were chosen to accurately identify sojourn times in the states. The proposed model enables the assessment of the probability of occurrence of individual states, especially the repair state, and its duration. Within the studied process, three operational states describing the machine's activity and two states assigned to renewal processes were distinguished, based on which a semi-Markov model was constructed, and the technical readiness was assessed.

However, the most significant achievement of this study is demonstrating that the assessment of the repair state's sojourn time is a mixture of different exponential distributions, each dependent on the transition to the next state. The precise identification of the repair state's sojourn time influences the identification of the readiness of the entire system. Neglecting this element can lead to erroneous conclusions and result in an unreliable assessment of the tested machine's readiness level, ultimately leading to a misjudgment of the entire system. Therefore, the authors' intention was to highlight the discrepancies in results – especially regarding the system's boundary stabilization – depending on the sequence of states executed. The presented method also has utilitarian value, allowing for inference regarding the timely completion of designated tasks by the machine and enabling the alignment of performed activities with the requirements of forestry management.

References

1. Duhovnik, A., Balabanova, Y., Lushpaeva, M., Gaiduk, A., & Nurullin, A. (2023). Analysis of promising methods for felling for forest care by multi-operation systems of machines. In E3S Web of Conferences (Vol. 390, p. 07043). EDP Sciences. https://doi.org/10.1051/e3sconf/202339007043
2. La Hera, P., Morales, D. O., & Mendoza-Trejo, O. (2021). A study case of Dynamic Motion Primitives as a motion planning method to automate the work of forestry cranes. Computers and Electronics in Agriculture, 183, https://doi.org/10.1016/j.compag.2021.106037.
3. Munis RA, Almeida RO, Camargo DA, da Silva RBG, Wojciechowski J, Simões D. Machine Learning Methods to Estimate Productivity of Harvesters: Mechanized Timber Harvesting in Brazil. Forests. 2022; 13(7):1068. https://doi.org/10.3390/f13071068
4. Söderberg, J., Wallerman, J., Almäng, A., Möller, J. J., & Willén, E. (2021). Operational prediction of forest attributes using standardised harvester data and airborne laser scanning data in Sweden. Scandinavian Journal of Forest Research, 36(4), 306–314. https://doi.org/10.1080/02827581.2021.1919751
5. Melander, L., Ritala, R. Separating the impact of work environment and machine operation on harvester performance. Eur J Forest Res 139, 1029–1043 (2020). https://doi.org/10.1007/s10342-020-01304-5
6. Dvořák, J., Kováč, J., & Krilek, J. (2020). Ergonomic Operational Working Aspects of Forest Machines. Cambridge Scholars Publishing.
7. Rozporządzenie Ministra Przedsiębiorczości i Technologii z dnia 30 października 2018 r. w sprawie warunków technicznych dozoru technicznego w zakresie eksploatacji, napraw i modernizacji urządzeń transportu bliskiego (Dz.U. 2018 poz. 2176)
8. Liski, E., Jounela, P., Korpunen, H., Sosa, A., Lindroos, O., & Jylhä, P. (2020). Modeling the productivity of mechanized CTL harvesting with statistical machine learning methods. International Journal of Forest Engineering, 31(3), 253-262. https://doi.org/10.1080/14942119.2020.1820750
9. Maktoubian, J., Taskhiri, M. S., & Turner, P. (2021). Intelligent predictive maintenance (Ipdm) in forestry: A review of challenges and opportunities. Forests, 12(11), https://doi.org/10.3390/f12111495.
10. Abbasi, R., Martinez, P., & Ahmad, R. (2022). The digitization of agricultural industry–a systematic literature review on agriculture 4.0. Smart Agricultural Technology, 2, https://doi.org/10.1016/j.atech.2022.100042.
11. Meyn, S. P., & Tweedie, R. L. (2012). Markov chains and stochastic stability. Springer Science & Business Media.
12. Davis, M. H. (2018). Markov models and optimization. Routledge. https://doi.org/10.1201/9780203748039
13. Howard, R. A. (2012). Dynamic probabilistic systems: Markov models. Courier Corporation.
14. Eriksson M. 2014. Productivity of harvesters and forwarders in CTL operations in northern Sweden based on large follow-up datasets. International Journal of Forest Engineering 25 (3). https://doi.org/10.1080/14942119.2014.974309
15. Dvořák J., Walczyk J., Natov P., Hošková P. (2015). Struktura czasu pracy harwesterów podczas pozyskania przygodnego. Structure of the operating time of the harvesters during casual logging. Sylwan 159 (4): 300−306.
16. Ziółkowski, J., Oszczypała, M., Lęgas, A., Konwerski, J., Małachowski, J. (2024). A method for calculating the technical readiness of aviation refuelling vehicles. Eksploatacja i Niezawodność – Maintenance and Reliability. https://doi.org/10.17531/ein/187888
17. Kubica, J., Ahmed, B., Muhammad, A., Aslam, M. U. (2024). Dynamic reliability calculation of random structures by conditional probability method. Eksploatacja i Niezawodność – Maintenance and Reliability, 26(2). https://doi.org/10.17531/ein/181133
18. Kozłowski, E., Borucka, A., Oleszczuk, P., Jałowiec, T. (2023). Evaluation of the maintenance system readiness using the semi-Markov model taking into account hidden factors. Eksploatacja i Niezawodność – Maintenance and Reliability, 25(4). https://doi.org/10.17531/ein/172857
19. Oszczypała, M., Ziółkowski, J., Małachowski, J. (2023). Semi-Markov approach for reliability modelling of light utility vehicles. Eksploatacja i Niezawodność – Maintenance and Reliability, 25(2), 1 25(2). https://doi.org/10.17531/ein/161859
20. Zhang, Q., Yang, L., Duan, J., Qin, J., Zhou, Y. (2024). Research on integrated scheduling of equipment predictive maintenance and production decision based on physical modeling approach. Eksploatacja i Niezawodność – Maintenance and Reliability, 26(1). https://doi.org/10.17531/ein/175409
21. Borucka A. Method of testing the readiness of means of transport with the use of semi-Markov processes, Transport 36 (1), 2021, https://doi.org/10.3846/transport.2021.14370
22. Borucka A, Kozłowski E, Parczewski R, Antosz K, Gil L, Pieniak D. Supply Sequence Modelling Using Hidden Markov Models. *Applied*

Sciences. 2023; 13(1):231. https://doi.org/10.3390/app13010231

23. Wang Z, Shangguan W, Peng C, Cai B. Similarity Based Remaining Useful Life Prediction for Lithium-ion Battery under Small Sample Situation Based on Data Augmentation. Eksploatacja i Niezawodność – Maintenance and Reliability. 2024;26(1). https://doi.org/10.17531/ein/175585
24. Murthy, D. P., Xie, M., & Jiang, R. (2004). Weibull models. John Wiley & Sons.
25. Ahmed Zohair D, Hafaifa A, Abdelhamid I, Abdellah K. Gas turbine reliability estimation to reduce the risk of failure occurrence with a comparative study between the two-parameter Weibull distribution and a new modified Weibull distribution. Diagnostyka. 2022;23(1):2022107. https://doi.org/10.29354/diag/146240
26. Thom, H. C. (1958). A note on the gamma distribution. Monthly weather review, 86(4), 117-122. https://doi.org/10.1175/1520-0493(1958)086%3C0117:ANOTGD%3E2.0.CO;2
27. Shorack, G. R., & Wellner, J. A. (2009). Empirical processes with applications to statistics. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898719017
28. Rausand, M., & Hoyland, A. (2003). System reliability theory: models, statistical methods, and applications (Vol. 396). John Wiley & Sons.
29. Almalki, S. J., & Nadarajah, S. (2014). Modifications of the Weibull distribution: A review. Reliability Engineering & System Safety, 124, 32-55. https://doi.org/10.1016/j.ress.2013.11.010
30. Kobayashi H, Mark BL, Turin W. Probability, random processes, and statistical analysis, Cambridge University Press 2011, https://doi.org/10.1017/CBO9780511977770
31. Phani, K. K. (1987). A new modified Weibull distribution function. Journal of the American Ceramic Society, 70(8), C-182. https://doi.org/10.1111/j.1151-2916.1987.tb05719.x
32. Mudholkar, G. S., & Kollia, G. D. (1994). Generalized Weibull family: a structural analysis. Communications in statistics-theory and methods, 23(4), 1149-1171. https://doi.org/10.1080/03610929408831309
33. Stacy, E. W., & Mihram, G. A. (1965). Parameter estimation for a generalized gamma distribution. Technometrics, 7(3), 349-358. https://doi.org/10.1080/00401706.1965.10490268