Article citation info:
SELECH J, ANDRZEJCZAK K. An aggregate criterion for selecting a distribution for times to failure of components of rail vehicles. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 102–111, http://dx.doi.org/10.17531/ein.2020.1.12.

Jarosław SELECH
Karol ANDRZEJCZAK

AN AGGREGATE CRITERION FOR SELECTING A DISTRIBUTION FOR TIMES TO FAILURE OF COMPONENTS OF RAIL VEHICLES

ZAGREGOWANE KRYTERIUM WYBORU ROZKŁADU CZASU DO USZKODZENIA ELEMENTÓW POJAZDÓW SZYNOWYCH*

This paper presents an aggregate method of selecting a theoretical cumulative distribution function (CDF) for an empirical CDF. The method was intended to identify the time of reliable operation of a renewable technical object by applying three criteria based on the following statistics: the modified Kolmogorov–Smirnov (MK-S) statistic, the mean absolute deviation of the theoretical CDF from the empirical CDF, and a statistic calculated on the basis of a log-likelihood function. The values of these statistics were used to rank eleven probability distributions. The data for which calculations were made concerned failures of the driver's cab lock recorded during five years of operation of a fleet of 45 trams. Before calculating the statistics, the empirical CDF of the examined component was determined using the Kaplan–Meier estimator, and then, using the method of Maximum Likelihood Estimation, the parameters of the analysed theoretical distributions were estimated. The theoretical distributions were then ranked according to the values obtained for each of the assumed criteria: the lower the value for a given criterion, the higher the ranking position, indicating a better fit according to that criterion. Then, based on the three rankings and on weights assigned to the individual criteria, an aggregate criterion (referred to as DESV) was implemented to select the best-fitting probability distribution. The method assumes that the lowest DESV value corresponds to the best-fitting theoretical distribution. In the case of the examined component, this was found to be the generalised gamma distribution. It is shown that if the final decision is based on the aggregate criterion, which takes into account the three criteria for goodness of fit, the reliability of the estimation of the time-to-failure distribution increases, and thus mistakes resulting from the use of only one of the criteria can be avoided.

Keywords: *time to failure, estimation of probability distribution, reliability of rail vehicles.*

W pracy przedstawiono zagregowaną metodę doboru dystrybuant hipotetycznych do dystrybuanty empirycznej. Metoda miała na celu identyfikację czasu niezawodnej pracy odnawialnego obiektu technicznego poprzez zastosowanie trzech kryteriów, w których użyto następujących statystyk: zmodyfikowanej statystyki Kołmogorowa-Smirnowa (MK-S), statystyki średniego odchylenia bezwzględnego dystrybuanty hipotetycznej od empirycznej oraz statystyki obliczanej na podstawie zlogarytmowanej funkcji wiarygodności. Wartości tych statystyk posłużyły do rangowania jedenastu rozkładów prawdopodobieństwa. Dane dla których dokonano obliczeń dotyczyły uszkodzeń zamka kabiny motorniczego jakie odnotowano w ciągu pięciu lat użytkowania floty 45 tramwajów. Przed obliczeniem statystyk wyznaczono dystrybuantę empiryczną badanego elementu przy pomocy estymatora Kaplana-Meiera, a następnie przy użyciu metody największej wiarygodności oszacowano parametry uwzględnionych w badaniach rozkładów hipotetycznych. Po wyznaczaniu parametrów nastąpiło rangowanie rozkładów hipotetycznych według wartości otrzymanych dla każdego z przyjętych kryteriów, im mniejsza wartość dla danego kryterium tym wyższa pozycja w rankingu, świadcząca o lepszej jakości dopasowania według danego kryterium. Po ustaleniu rankingu według kryteriów zgodności, każdemu z kryteriów zgodności dopasowania dystrybuant modelowych do empirycznej nadano wagi. Następnie na podstawie uzyskanych trzech rankingów oraz wag nadanych poszczególnym kryteriom zgodności wyznaczana jest zagregowana miara zgodności (oznaczona DESV), która służy do wyznaczania najlepszego rozkładu prawdopodobieństwa. W prezentowanej metodzie przyjęto, że najmniejsza wartość DESV wyznacza najlepiej dopasowany rozkład hipotetyczny. W przypadku badanego elementu rozkładem tym okazał się uogólniony rozkład gamma. Pokazano, że na podstawie zagregowanego kryterium uwzględniającego trzy statystyki zgodności dopasowania zwiększa się wiarygodność estymacji rozkładu czasu pracy do uszkodzenia, unikając tym samym błędów jakie można popełnić uzależniając się tylko od jednej z nich.

Słowa kluczowe: *czas do uszkodzenia, estymacja rozkładu prawdopodobieństwa, niezawodność pojazdów szynowych.*

1. Introduction

In traditional methods of estimating the parameters of the time-to-failure distribution of a technical object or its components, a specific distribution class is assumed *a priori*. The purpose of this article is to present the results of a procedure to identify the best-fitting probability distribution model for the time to failure of a renewable technical object using an aggregate criterion. The research concerns components of currently operated rail vehicles of a uniform type that belong to a fleet maintained by the operator. Empirical data obtained during the operation of the vehicles are incomplete, since the vehicles were operational at the end of the data acquisition period. Thus, the authors did not have complete data on the times to failure of all components of the analysed vehicles. Therefore, it was necessary to use statistical methods taking account of censored data. Given a suitably prepared database of repairs to vehicles in the fleet, it is relatively easy to de-

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

Table 1. Density functions and parameters to be estimated

Distribution model	Probability density function	Distribution parameters
Exponential	$f(t;\lambda)=\lambda e^{-\lambda t},\ t\ge 0,\lambda>0$	$\frac{1}{\lambda}$ – scale parameter
Two-parameter exponential	$f(t;\lambda,\gamma)=\lambda e^{-\lambda(t-\gamma)},\ t\ge\gamma,\lambda>0$	$\frac{1}{\lambda}$ – scale parameter γ – location parameter
Normal	$f(t;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2},\ t\in\mathbb{R},\mu\in\mathbb{R},\sigma>0$	μ – expected value σ – standard deviation
Lognormal	$f(t;\mu',\sigma')=\frac{1}{t\cdot\sigma'\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln(t)-\mu'}{\sigma'}\right)^2},\ t>0,\mu'\in\mathbb{R},\sigma'>0$	μ' – expected value of normally distributed ln T σ' – standard deviation of ln T
Two-parameter Weibull	$f(t;\beta,\eta)=\frac{\beta}{\eta}\cdot\left(\frac{t}{\eta}\right)^{\beta-1}\cdot e^{-\left(\frac{t}{\eta}\right)^{\beta}},\ t\ge 0,\beta>0,\eta>0$	η – scale parameter β – shape parameter
Three-parameter Weibull	$f(t;\beta,\eta,\gamma)=\frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1}e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}},\ t\ge\gamma,\beta>0,\eta>0,\gamma\in\mathbb{R}$	η – scale parameter β – shape parameter γ – location parameter
Gamma	$f(t;\mu,\kappa)=\frac{\exp\left(\kappa(\ln(t)-\mu)-\exp\left(\ln(t)-\mu\right)\right)}{t\Gamma(\kappa)};\ t>0,\mu\in\mathbb{R},\kappa>0$	e^{μ} – scale parameter κ – shape parameter
Generalised gamma	$f(t;\theta,\beta,\kappa)=\frac{\beta}{\Gamma(\kappa)\cdot\theta}\cdot\left(\frac{t}{\theta}\right)^{\kappa\beta-1}\cdot e^{-\left(\frac{t}{\theta}\right)^{\beta}},\ \ \theta>0,\beta>0,\kappa>0$ Reparameterisation: $\mu=\ln(\theta)+\frac{1}{\beta}\ln\left(\frac{1}{2}\right);\ \ \sigma=\frac{1}{\beta\sqrt{\kappa}};\ \ =\frac{1}{\sqrt{\kappa}}$ $f(t;\mu,\sigma,\lambda)=\begin{cases}\frac{\lvert\lambda\rvert}{\sigma\cdot t}\cdot\frac{1}{\Gamma\left(\frac{1}{2}\right)}\cdot\exp\left[\frac{\lambda\frac{\ln(t)-\mu}{\sigma}+\ln\left(\frac{1}{\lambda^2}\right)-\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)}{\lambda^2}\right] & if\ \lambda>0\\ \frac{1}{t\cdot\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(t)-\mu}{\sigma}\right)^2\right) & if\ \lambda=0\end{cases}$ $t\ge 0,\mu\in\mathbb{R},\lambda\ge 0,\sigma>0,$	θ – scale parameter β – shape parameter κ – shape parameter
Logistic	$f(t;\mu,\sigma)=\frac{\exp\left(\frac{t-\mu}{\sigma}\right)}{\sigma\left(1+\exp\left(\frac{t-\mu}{\sigma}\right)\right)^2},t\in\mathbb{R},\ \mu\in\mathbb{R},\sigma>0$	σ – scale parameter μ – location parameter
Loglogistic	$f(t;\mu,\sigma)=\frac{\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)}{\sigma t\left(1+\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)\right)^2},\ t>0,\ \mu\in\mathbb{R},\sigma>0$	μ – scale parameter σ – shape parameter
Gumbel	$f(t;\mu,\sigma)=\frac{1}{\sigma}\exp\left(\left(\frac{t-\mu}{\sigma}\right)-\exp\left(\frac{t-\mu}{\sigma}\right)\right),\ \sigma>0$	μ – location parameter σ – scale parameter

termine basic reliability characteristics of failed components [25]. However, the selection of a good criterion for the fit of a distribution of times to failure of the components becomes an issue. This problem is the subject of the research presented in this paper, which concerns the use of an aggregate criterion for determining the best-fitting time-to-failure distributions for selected components of a rail vehicle [33]. The research results are presented in the form of a ranking of the fit of selected families of distributions based on the aggregate criterion.

In the study of technical objects, different probability distribution families are used as models of time to failure [17]. The most commonly used distributions in Life Data Analysis (LDA) are the normal, exponential and Weibull distributions [19, 10]. In this study, apart from the aforementioned distributions, the authors also verified the possibility of using other, less common distributions, whose goodness of fit to the empirical data proved superior in many cases to the more common distributions. These are the lognormal, gamma, generalised gamma, logistic, loglogistic and Gumbel distributions [22]. The density functions of these distributions and their parameters are listed in Table 1. In the case of the generalised gamma distribution, for easier parameter estimation, the density function is also given in reparameterised form [20].

The parameters of these distributions can be estimated using analytical, numerical and graphical methods [16, 26, 29]. The most commonly used methods include the method of moments, Maximum Likelihood Estimation, the least squares method, the method of probability plotting, and the probability plot correlation coefficient (PPCC) method [1, 38, 32]. In engineering practice, the most commonly used are numerical and graphic methods executed with specialised IT tools [12, 39]. Based on the operational data and selection of the estimation method, parameters (shape, scale, location) are estimated for selected families of probability distributions [28, 37]. Having estimated various distributions, it is possible to indicate which of them is the best fitted to the empirical data in the sense of the lowest sum of squares of deviations.

The proposed methodology for identifying the time to failure of a selected vehicle component uses all available data on times (mileages) between failures of the component in all vehicles of the analysed fleet. This includes the case where the component was operational at the time when data acquisition was ended; the time to failure of such a component is said to be right-censored. A method of preparing statistical data based on the operational database has been developed in the articles [3, 2].

Instead of the traditional single-criterion selection of the best-fitting family of probability distributions, the authors propose to use an aggregate criterion that includes three measures of the fit of theoretical distributions. This criterion takes into account a ranking of the fit of individual probabilistic models to the empirical data, including right-censored operational data for the vehicle fleet.

In the aggregate method, the choice of a distribution is preceded by a ranking of distributions for three goodness-of-fit criteria. The parameters of selected distribution families were estimated using the Weibull++ Distribution Wizard module, which – after performing the appropriate calculations – ranks the distributions starting from that with the highest goodness of fit. However, before the fit of the distributions is examined, the CDF or reliability function of the empirical distribution is determined by the Kaplan–Meier method, and then the parameters of the theoretical distributions are determined by the method of Maximum Likelihood Estimation (MLE).

The next step is to determine the statistics of the goodness of fit of the theoretical CDFs to the empirical CDF, denoted F_n. On this basis, a ranking is made of eleven distributions, listed in Table 1, that are used in the survival analysis [18, 30]. Provided that the assumptions are met, the rankings of distributions by goodness of fit are compiled independently according to three criteria, using the modified Kolmogorov–Smirnov (MK-S) statistic, the statistic of the mean absolute deviation of the theoretical CDF from the empirical CDF, and the value of the log-likelihood function [23].

The final ranking of distributions is based on the rankings obtained using these three criteria, taking into account the weights assigned to each of them. After assigning weights to the criteria, the final Distribution Estimation Values (DESV) are calculated, indicating the best-fitting distribution according to the aggregate criterion. The scheme of successive calculation steps in the aggregate method of ranking distributions is shown in Fig. 1.

According to this scheme, in the first step, based on the obtained data and analysing the length of the observation time (right-censored), the survival function parameters were estimated with the Kaplan–Meier estimator and an empirical CDF was determined [7]. Then, to determine the parameters of the eleven theoretical distributions listed in Table 1, the method of Maximum Likelihood Estimation was used [15, 11].

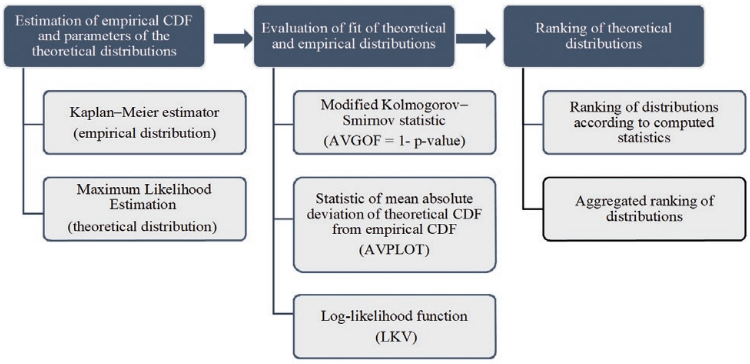

Fig. 1. Aggregate criterion for ranking distributions

In the second step, for each of the eleven distributions, the goodness-of-fit statistics are used to test the null hypothesis:

$$H_0 : T \sim F \tag{1}$$

stating that the time to failure T of the analysed vehicle component has a probability distribution with the CDF F with the estimated parameters. This evaluation is based on a random sample $T_1, T_2, \ldots, T_n$ concerning times to failure of the component. In this paper, the times to failure of the examined component are expressed in terms of kilometres travelled, as in the paper [2].

2. Criteria for ranking theoretical distributions

Among the applied goodness-of-fit criteria, a particular role is played by the modified Kolmogorov–Smirnov statistic (AVGOF, average goodness of fit), which evaluates the statistical difference between the values of the empirical and theoretical CDFs. The particular role of this statistic results from the fact that it is highly sensitive to local deviations. In addition, it can be used even with a small amount of data and with unknown parameters of the theoretical distribution. The use of the MK-S statistic is therefore necessary when the parameters of the tested distributions need to be estimated.

Because the distributions of MK-S statistics depend on a theoretical distribution family whose parameters are estimated, a critical value, at which the null hypothesis is rejected, is determined for each distribution [30]. Analytical determination of the critical value is often difficult or even impossible, and hence it is obtained using the Monte Carlo method [6, 18].

The MK-S statistic used to test the fit of a theoretical distribution to the empirical distribution uses the statistic D_{max}, defined as the maximum of the absolute difference between the value of the empirical CDF $F_n(t)$ and a matched theoretical CDF $F(t)$, and given by the formula [18]:

$$D_{max} = \max_{1 \le i \le n} \left| F_n(t_i) - F(t_i) \right| \tag{2}$$

where:

D_{max} is the value of the statistic;
n is the sample size;
$F_n(t_i)$ is the value of the empirical CDF;
$F(t_i)$ is the value of the theoretical CDF.

The critical value D_{CRIT} in the modified Kolmogorov–Smirnov statistic is determined by the Monte Carlo method, as already mentioned, due to the difficulty of the calculations.

The MK-S statistic is used to determine the probability of rejection of the null hypothesis, i.e. the probability of the event $D_{CRIT} < D_{max}$. Hence, in the case of the first criterion, the basis for ordering theoretical distributions is the probability:

$$\mathrm{P}\left(D_{CRIT} < D_{max}\right) \tag{3}$$

The higher the value of the statistic D_{max}, the more significant is the difference between the theoretical distribution defined by the CDF F and the empirical distribution with the CDF F_n. Because the critical value D_{CRIT} is determined by the Monte Carlo method through m-tuple generation of n time-to-failure values $t_{s1}, t_{s2}, \dots t_{sn}$, for which simulation CDFs $F_s(t_{si}), s = 1,2,\dots,m$ are created, and maximum differences with the values of the theoretical CDF are determined for each of these:

$$d_{max,s} = \max_{1 \le s \le m} \left| F_s(t_{si}) - F(t_{si}) \right|, s = 1,2,\dots,m \tag{4}$$

the critical value D_{CRIT} is estimated as the arithmetic mean d_{CRIT} defined by the formula [6]:

$$\widehat{D_{CRIT}} = d_{CRIT} = \frac{1}{m}\sum_{s=1}^{m} d_{max,s} \tag{5}$$

Finally, in the MK-S criterion for the goodness of fit of distributions we assume:

$$\mathrm{AVGOF}\left(F_n, F\right) = 100 \cdot \mathrm{P}\left(d_{CRIT} < D_{max}\right) \tag{6}$$

Large values of AVGOF, close to 100, indicate that there is a significant difference between the theoretical distribution and the empirical data. Hence, the lower the value of the statistic AVGOF, the better the fit of the theoretical distribution.

In the case of the second goodness-of-fit criterion, the mean absolute deviation of the theoretical CDF from the empirical CDF is examined, and the statistic used to assess goodness of fit, denoted AVPLOT (average plot fit), is determined according to the formula:

$$\mathrm{AVPLOT}\left(F_n, F\right) = 100\frac{1}{n}\sum_{i=1}^{n} \left| F_n(t_i) - F(t_i) \right| \tag{7}$$

where:

n is sample size;
$F_n(t_i)$ are values of the empirical CDF;
$F(t_i)$ are values of the theoretical CDF.

This criterion, unlike the MK-S criterion, is not sensitive to local deviations, but takes into account the global difference of the distributions and is a good complement to the MK-S criterion.

For the third criterion for testing the fit of distributions, the likelihood function (LKV, Likelihood Value Test) was used as a measure of the fit of a probabilistic model to empirical data. The log value of the likelihood function (LKV) is calculated for empirical data [27, 14]. The likelihood function L depends on the random sample $T_1, T_2, \dots, T_n$ and on parameters θ_j for which it takes maximum values. The general form of the likelihood function is given by the formula [33, 30]:

$$L\left(\theta_1, \theta_2, \dots, \theta_k \middle| T_1, T_2, \dots, T_n\right) = \prod_{i=1}^{n} f\left(T_i; \theta_1, \theta_2, \dots, \theta_k\right) \tag{8}$$

where:

n is the number of failed components;
k is the number of parameters;
$\theta_j, j = 1,2,\dots,k$ is the *j*-th parameter of the distribution;
$T_i, i = 1,2,\dots,n$ is the time to failure of the *i*-th component.

In the case under consideration, the function was expanded to include factors taking account of right-censored data. The log-likelihood function is the sum of logarithms of probability density for particular lifetimes of the analysed component [18, 30]:

$$\Lambda\left(\theta_1, \theta_2, \dots, \theta_k\right) = \ln L\left(\theta_1, \theta_2, \dots, \theta_k \middle| T_1, T_2, \dots, T_n\right) = \sum_{i=1}^{n} \ln f\left(T_i; \theta_1, \theta_2, \dots, \theta_k\right) \tag{9}$$

where:

L is the likelihood function;
n is the number of failed components;
is the *j*-th parameter of the distribution;
T_i, $i = 1,2,\dots,n$ is the time to failure of the i-th component.

Values of estimators of the unknown parameters $\theta_1, \theta_2, \dots, \theta_k$ are determined by maximising the log-likelihood function $\Lambda\left(\theta_1, \theta_2, \dots, \theta_k\right)$. A necessary condition for the existence of an extremum of this function is that all of its partial derivatives take the value 0.

To determine the estimators of the unknown parameters, partial derivatives $\frac{\partial \Lambda\left(\theta_1, \theta_2, \dots, \theta_k\right)}{\partial \theta_j}$ of the function Λ are determined with respect to the parameters θ_j, $j = 1,2,\dots,k$. To estimate the parameters, each partial derivative should be equated to zero and k equations should be solved:

$$\begin{gathered} \frac{\partial \Lambda\left(\theta_1, \theta_2, \dots, \theta_k\right)}{\partial \theta_1} = 0 \\ \cdots\cdots\cdots \\ \frac{\partial \Lambda\left(\theta_1, \theta_2, \dots, \theta_k\right)}{\partial \theta_k} = 0 \end{gathered} \tag{10}$$

In the last step, on the basis of each of the three goodness-of-fit criteria for all 11 distributions, ranks are assigned from the best-fitting to the worst-fitting theoretical distribution. Thus, the theoretical distributions are ordered separately for each criterion by assigning them successive natural numbers. Finally, based on the three rankings obtained and the weights assigned to the individual criteria, the aggregate criterion DESV is determined. This measure, for the *i*-th theoretical CDF (F_i), is given by formula (11):

$$DESV(F_i)=RAVGOF(F_i)\cdot WAVGOF+RAVPLOT(F_i)\cdot WAVPLOT+ RLKV(F_i)\cdot WLKV \quad (11)$$

where:

- $RAVGOF(F_i)$ denotes the rank of the distribution F_i by the $AVGOF$ criterion;
- $RAVPLOT(F_i)$ denotes the rank of the distribution F_i by the $AVPLOT$ criterion;
- $RLKV(F_i)$ denotes the rank of the distribution F_i by the LKV criterion;
- $WAVGOF$ denotes the weight of the $AVGOF$ criterion;
- $WAVPLOT$ denotes the weight of the $AVPLOT$ criterion;
- $WLKV$ denotes the weight of the LKV criterion.

The aggregate criterion $DESV$ is therefore a weighted average of the individual ranks of theoretical distributions. After calculating the DESV value for the particular theoretical distributions, their final ranking is determined. The distribution with the lowest DESV value is identified as the best-fitting according to the aggregate criterion, and is assigned the number 1 in the ranking. The aggregate criterion is used to make the final selection of the distribution that best fits the empirical data among the theoretical distributions considered.

3. Subject of study

The aggregate criterion for ranking distributions of times to failure of selected vehicle components was applied based on operational data from a fleet of 45 urban rail vehicles of the same type, namely five-section low-floor Tramino S105P trams with total weight 42.5 tonnes and length approximately 32 metres. These are articulated, single-compartment vehicles. The tram can carry a maximum of 229 passengers, including 48 seated. The operational data covered the initial five years of use of the fleet, including two years covered by the warranty and three subsequent years under a maintenance contract [31, 9]. All trams were used in similar operating conditions, i.e. the same track infrastructure, similar daily and annual times of travel, and the same schedule and scope of (preventive) maintenance.

Fig. 2. Tramino S105P tram

From the database of failures in trams of the fleet under investigation, the lock of the driver's cab door was selected for testing of the time-to-failure distribution. This component failed 54 times during the first five years of operation, and generated 0.52% of all corrective maintenance costs [5]. The lock is mounted on the door between the passenger space and the driver's cab. To open the driver's cab door from the outside, the lock has to be opened mechanically with a special key. It was the bolting part of the lock that failed, becoming blocked and thus preventing the driver from opening the door and entering the cab. Depending on where the failure occurred, it was necessary to call the emergency maintenance service or to open the door using force, damaging the strike plate structure. On each such occasion the damaged lock was replaced with a new one. The cause of failure of the lock was excessive wear of the internal mechanism responsible for bolt extension, caused by a poorly selected construction material, as a result of which the lock stuck and sometimes prevented removal of the inserted key. A photograph of the lock is shown in Fig. 3.

Fig. 3. Driver's cab door lock

4. Empirical data

The process of tram operation is a valuable source of information serving to assess the required reliability parameters and to forecast maintenance costs. Operational information should be taken to include all data on events occurring during the operation and maintenance of trams [13]. These data play a key role in the planning and day-to-day management of vehicle fleet operation and maintenance, as well as in improving vehicle technology and construction [4, 35]. Operational information plays a particularly important role for operating companies, as it enables the proper planning of costs of operation, inspections and repairs, as well as assessment of the use of the means of transport [24, 34].

Before proceeding to the estimation of the parameters of probabilistic models of times to failure of selected vehicle components, the operational data should be appropriately prepared. For the investigated fleet of trams, operational data regarding individual vehicle components is right-censored of type I, which means that for a fixed period of use of the tram fleet the lock

Table 2. Right-censored times to failure of the lock in 5 years of operation

Dist. travelled [km]	F/S	Dist. travelled [km]	F/S	Dist. travelled [km]	F/S	Dist. travelled [km]	F/S
174,124	F	256,382	F	114,128	S	67,733	F
196,837	S	144,819	S	135,078	F	300,557	S
317,275	S	223,684	F	136,600	F	103,378	F
292,525	F	46,217	F	97,832	S	177,506	F
112,431	S	43,897	S	377,101	S	23,153	S
196,218	F	155,522	F	93,585	F	242,544	F
1,910	F	201,423	S	238,103	S	89,047	S
93,529	S	119,376	F	285,538	F	125,785	F
334,484	S	198,190	S	43,117	S	58,407	F
366,935	F	368,449	S	221,226	F	117,646	S
28,826	S	340,330	F	117,701	S	202,396	F
191,367	F	58,964	S	28,934	F	127,143	S
21,117	F	193,641	F	135,673	F	287,695	F
135,831	S	155,920	S	155,828	S	53,863	S
348,956	F	206,246	F	92,594	F	174,580	F
38,020	S	144,352	S	197,981	S	139,571	S
188,493	F	371,800	S	148,840	F	210,775	F
70,534	F	22,482	F	27,858	F	102,038	S
102,343	S	139,974	F	107,491	F	131,537	F
340,236	F	39,840	F	52,280	S	126,738	F
52,022	S	127,333	F	250,370	S	83,497	S
115,592	F	21,021	S	282,989	F	176,928	F
79,071	F	376,601	S	77,834	S	81,021	F
72,135	F	354,513	S	86,028	F	103,807	S
105,552	S	203,105	F	226,082	S	-	-

Table 3. Estimated parameters of tested distributions

1P-Exponential	**2P-Exponential**	**Normal**	**Lognormal**
$\hat{\lambda}$ = 3.413E-06	$\hat{\lambda}$ = 3.937E-06	$\hat{\mu}$ = 218,279.5	$\hat{\mu}'$ = 12.184
	$\hat{\gamma}$ = 21.117	$\hat{\sigma}$ = 115,461.8	$\hat{\sigma}'$ = 0.819
2P-Weibull	**3P-Weibull**	**Gamma**	**G-Gamma**
$\hat{\beta}$ = 1.745	$\hat{\beta}$ = 1.885	$\hat{\mu}$ = 11.53	$\hat{\mu}$ = 12.415
$\hat{\eta}$ = 255,316.9	$\hat{\eta}$ = 266,209.6	$\hat{\kappa}$ = 2.307	$\hat{\sigma}$ = 0.605
	$\hat{\gamma}$ = −10,300.12		$\hat{\lambda}$ = 0.857
Logistic	**Loglogistic**	**Gumbel**	
$\hat{\mu}$ = 211,755.9	$\hat{\mu}$ = 12.198	$\hat{\mu}$ = 274,770.6	
$\hat{\sigma}$ = 68,491.7	$\hat{\sigma}$ = 0.447	$\hat{\sigma}$ = 104,341.5	

F – failure, S – survival

failed and was replaced only in some vehicles, while in some vehicles it was replaced multiple times. Because the research concerns vehicles that are operated intensively, times to failure of individual components are expressed in kilometres. The time at which each vehicle comes into operation is known, and tram mileages at which components fail are recorded [33, 2]. The mileage of trams is used to determine the mileage of components at failure. The method of determining the mileage of failed vehicle components is presented in the paper [3].

Suitably prepared data are summarised in Table 2. They contain the exact time to failure of the tested component (the driver's cab lock) from the fleet of 45 trams under observation, expressed in kilometres and marked as F (failure), and the survival time of other locks that did not fail, marked as S (survival), also expressed as a number of kilometres travelled until the observations ended. At the time of the end of observations, the locks in all 45 vehicles were functional, although many of them had been replaced due to failure. Because the main reason for the replacement of locks is failure in the opening mechanism, all failures of this type were classified as mechanical failure.

Based on the data in Table 2, parameters were estimated for 11 theoretical distributions. The results of estimation for all examined distributions are given in tabular form (Table 3).

5. Identification of the best-fitting probability distribution

To select the best-fitting theoretical distribution out of the 11 considered, the aggregate ranking criterion described in section 2 was used. In determining the ranking of distributions, first the parameters of the theoretical distributions were estimated, and then the distributions were ranked based on the three criteria described. The results of this ranking procedure are summarised in Table 4.

The first column contains the name of the probability distribution. The second contains the values of the Kolmogorov–Smirnov AVGOF statistic – the probability of rejection of the working hypothesis for the MK-S statistic. The third column (AVPLOT) gives the mean absolute deviation of the theoretical CDF from the empirical CDF. The fourth column (LKV) gives the measures of goodness of fit determined using the log-likelihood criterion [8, 36, 21].

Table 4. Results of individual statistics for the data in Table 1

Distribution	AVGOF	AVPLOT	LKV
1P-Exponential	80.740	7.599	-720.16
2P-Exponential	55.253	5.666	-712.58
Normal	30.946	4.178	-715.65
Lognormal	14.276	2.790	-711.82
2P-Weibull	1.034	1.709	-709.71
3P-Weibull	2.396	1.870	-710.23
Gamma	0.045	1.585	-709.78
G-Gamma	**0.289**	**1.589**	**-709.66**
Logistic	36.386	3.730	-717.34
Loglogistic	1.716	1.768	-710.49
Gumbel	84.250	6.355	-723.73

After calculating the goodness-of-fit statistics for the three criteria and ranking the probability distributions, the next step was to assign weights to the criteria. In this study, the default values of weights selected by the software manufacturer were used. These are determined on the basis of engineering practice, resulting from many analyses conducted in industrial applications. Using the weights assigned to each criterion, the weighted average was calculated for the ranks obtained using the individual criteria. Finally, using the described DESV aggregate criterion, the final ranking of the eleven theoretical distributions was obtained. Weibull++ software was used for the estimation of parameters of the theoretical distributions and for constructing their rankings. For the analysed data, the following weights were assigned to the criteria: 40 for AVGOF, 10 for AVPLOT and 50 for LKV. After calculating the DESV value, the final ranking of distributions was determined (Table 5). The distribution with the lowest DESV value was identified as the best-fitting according to the aggregate criterion, and was assigned number 1 in the ranking. As shown in Table 5, the lowest value of the DESV statistic was obtained for the generalised gamma distribution. It was calculated from formula (11) as follows:

$$\mathrm{DESV} = \left(2\times 40\right)+\left(2\times 10\right)+\left(1\times 50\right)=150 \tag{12}$$

Thus, for the data contained in Table 2 regarding lock failures during five years of operation of the tram fleet, using the developed aggregate criterion, the generalised gamma distribution was identified as the best-fitting. This is reflected in the last column of Table 5.

Table 5. Weighted average values and ranking of distributions

Distribution	AVGOF	AVPLOT	LKV	DESV	Ranking
1P-Exponential	10	11	10	1010	9
2P-Exponential	9	9	7	800	7
Normal	7	8	8	760	6
Lognormal	6	6	6	600	5
2P-Weibull	3	3	2	250	3
3P-Weibull	5	5	4	450	4
Gamma	1	1	3	200	2
G-Gamma	**2**	**2**	**1**	**150**	**1**
Logistic	8	7	9	840	8
Loglogistic	4	4	5	450	4
Gumbel	11	10	11	1090	10

The estimated parameters μ, σ, λ for the reparameterised form of this distribution took the following values: $\hat{\mu}$=12.415; $\hat{\sigma}$=0.6058; $\hat{\lambda}$=0.8572. The calculated rate of failure of the lock was 0.000000617/km, and the average time to failure was 229,623 km.

To illustrate how the selected distribution matches the data, in Fig. 4 the data are presented on a probability plot of the generalised gamma distribution. The following figures show the reliability function (Fig. 5), the probability density function (Fig. 6) and a histogram of numbers of failures (Fig. 7).

In Fig. 4 the blue line represents the modelled probability of failure according to the generalised gamma distribution, and the red lines mark a two-sided 95% confidence interval. The reliability function graph (Fig. 5) shows the change in the reliability value over time, expressed as distance travelled in kilometres, indicating the trend in the behaviour of the tested component in terms of failures. The graph of the failure probability density function provides a visualisation of the distribution of data over time (Fig. 6). The histogram (Fig. 7) shows that a relatively large proportion of the failures occurred between 50,000 and 200,000 km.

The graphical presentation of the estimated functional characteristics (reliability, probability density) and the histogram of numbers of failures can be used to determine more easily the failure mode. This information is important when forecasting failures and determining the future cost of corrective maintenance resulting from them.

The presented analysis of the time to failure of the driver's cab lock shows that the best-fitting distribution, according to the aggregate criterion, is the generalised gamma distribution. It should also be noted that with successive failures, the aggregate method may indicate a different distribution as the best-fitting, because new data, especially if the quantity is large relative to that previously analysed, may follow a different model. In this situation, analysis of the plot of the probability distribution is very useful for pre-evaluating the fit of a selected theoretical model to the appropriate case.

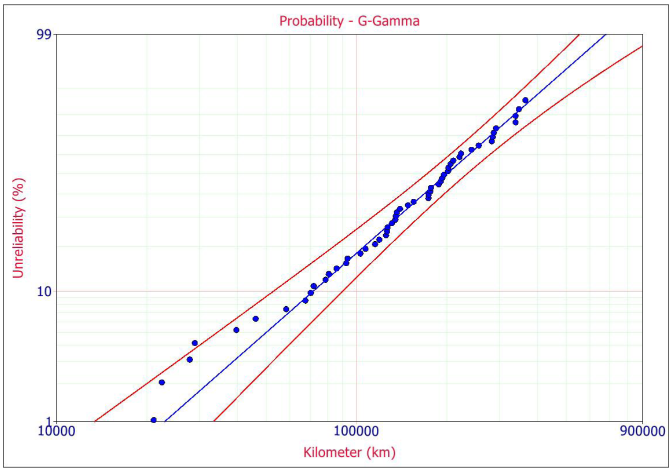

Fig. 4. Presentation of data on a probability plot of the generalised gamma distribution

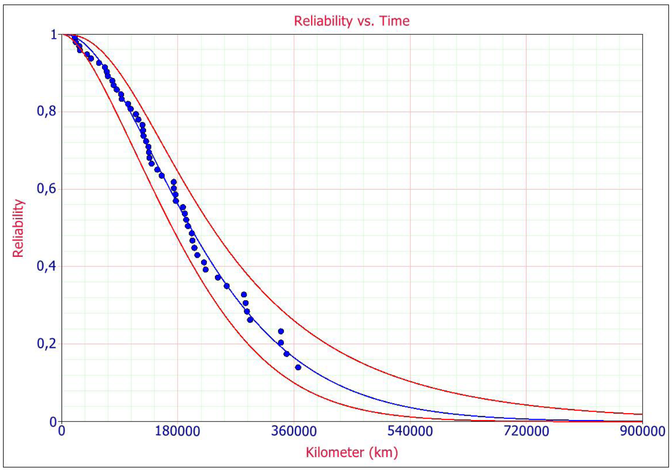

Fig. 5. Reliability function

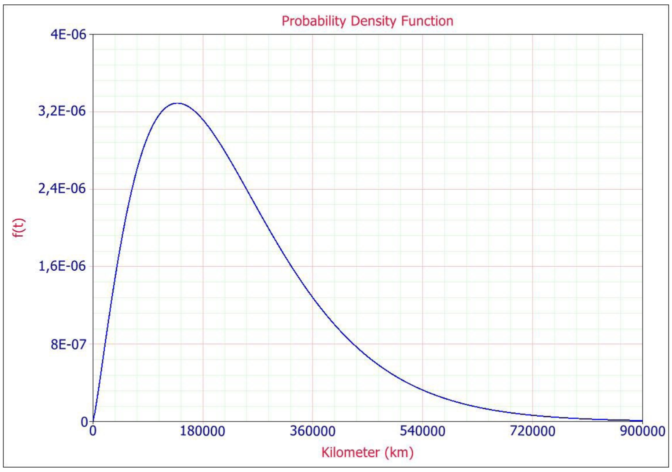

Fig. 6. Probability density function

When analysing failure data using an aggregate method, it has to be remembered that sometimes none of the statistical distribution models match the analysed data. In this case, the best of the worst solutions is obtained, which may poorly fit the data. In other cases, in which many models may be well matched to the empirical data, statistics alone are not enough; in such cases knowledge of the failure mechanism can be invaluable when selecting the most appropriate theoretical model. It is important to remember that, although the aggregate method applied to small samples will also rank selected

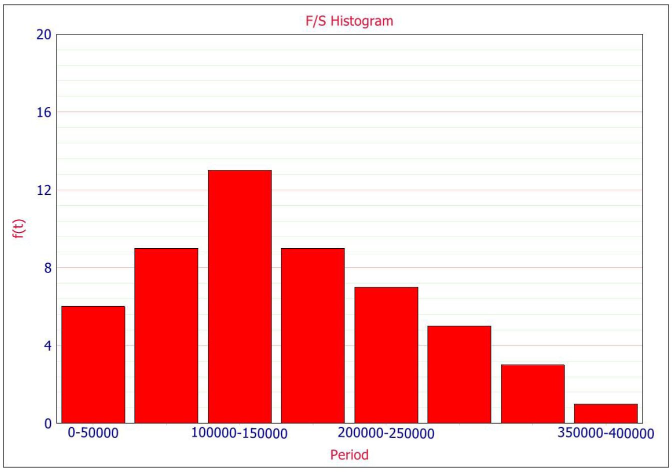

Fig. 7. Histogram

probability distributions depending on the number of parameters in a particular theoretical distribution, using it in such cases comes with a high level of uncertainty, and it is only recommended for use with larger data sets.

It should also be borne in mind that the two-parameter exponential distribution, the three-parameter Weibull distribution and the generalised gamma distribution contain a location parameter, a change in which causes a shift of the CDF and probability distribution function without changing their shape. On the other hand, the generalised gamma distribution is a complex model that can easily mimic many other distributions, and therefore often seems to be the best fitted to the analysed data.

When analysing data on probability plots, it can often be stated that they reflect more than one type of failure (e.g. fatigue, operational, construction, technological, etc.). In this case, all distributions ordered according to the aggregate selection method may turn out to be mismatched, because the developed method can only be used for a homogeneous type of failure of the examined component. In such situations, it is advisable to consider the possibility of using a mixture of distributions, e.g. a combination of two Weibull distributions.

6. Conclusions

The results obtained constitute an important argument for the possibility of using the proposed aggregate method of selecting a theoretical distribution for empirical data. By taking into account three criteria for assessing the accuracy of the fit, mistakes resulting from the use of only one of them can be avoided.

The use of only one criterion defining the quality of the fit of a theoretical to an empirical CDF may often prove insufficient, as it depends on many variables: mainly on the quantity of data and whether the data are full or censored, but primarily on the type of failure.

The aggregate method of identifying a theoretical distribution, taking into account three criteria, is a general method and has wide application, provided that the appropriate conditions are met: the number of observations must be large enough, and should contain accurate data on times to failure or to the end of observations. The modified K-S statistic (AVGOV) is sensitive to local deviations. On the other hand, the mean absolute deviation of the theoretical CDF from the empirical CDF (AVPLOT) is not so sensitive to local deviations; it takes into account the global difference of distributions, and is a good complement to the MK-S criterion. For the third criterion, the logarithm of the likelihood function (LKV), the size of the sample is important, because for small samples the value obtained may be strongly biased.

The benefits resulting from the correct selection of a random variable distribution for the time to failure of a renewable technical object (a rail vehicle) are significant, among others due to the costs generated by failing to utilise fully the potential lifetime of the component, as well as losses resulting from unplanned corrective maintenance and vehicle downtime.

The research was financed from the following funds: PUT 04/43/SBAD/0111 and 05/51/SBAD/3584.

References

1. Abernethy R B. The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Survivability, Risk, Cost, and Warranty Claims (Fifth ed.), Florida, 2010.
2. Andrzejczak K, Selech J. Flexible Prediction of the Vehicle Component Damage. Transport Means 2018: Proceedings of the 22nd International Scientific Conference, Trakai, Lithuania, Part II, 2018; 987-990.
3. Andrzejczak K, Selech J. Generalised Gamma Distribution in the Corrective Maintenance Prediction of Homogeneous Vehicles. In: Kabashkin I, Yatskiv (Jackiva) I, Prentkovskis O. (eds) Reliability and Statistics in Transportation and Communication. RelStat 2018. Lecture Notes in Networks and Systems. Springer, Cham 2018; 68.4. Andrzejczak K, Selech J. Investigating the trends of average costs of corrective maintenance of public transport vehicles. Journal of KONBiN 2017; 41: 207-226, https://doi.org/10.1515/jok-2017-0011.
5. Andrzejczak K, Selech J. Quantile analysis of the operating costs of the public transport fleet. Transport Problems, 2017; 12 (3): 103-111.
6. Andrzejczak K. Statystyka elementarna z wykorzystaniem systemu Statgraphics [Elementary statistics using the Statgraphics system], Wyd. Politechniki Poznańskiej, Poznań 1997.
7. Bartnik G, Pieniak D, Niewczas A M, Marciniak A. Probabilistic model for flexural strength of dental composites used in modelling reliability of the "tooth-dental composite" system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (1): 136-141, https://doi.org/10.17531/ein.2016.1.18.
8. Bavuso S J. Aerospace Applications of Weibull and Monte Carlo Simulation with Importance Sampling, IEEE, Annual Reliability and Maintainability Symposium, Proc. 1997.
9. Dolce J E. Analytical Fleet Maintenance Management, SAE International, SUA, 1994.
10. Elmahdy E E. Modelling Reliability Data with Finite Weibull or Lognormal Mixture Distributions. Appl. Math. Inf. Sci. 2017; 11 (9), 1081-1089, https://doi.org/10.18576/amis/110414.
11. Ferreira L A, Silva J L. Parameter estimation for Weibull distribution with right censored data using EM algorithm. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (2): 310-315, https://doi.org/10.17531/ein.2017.2.20.
12. Fuc P, Rymaniak L, Ziolkowski A. The correlation of distribution of PM number emitted under actual conditions of operation by PC and HDV vehicles, WIT Transactions on Ecology and the Environment. WIT Press, 2013; 174: 207.13. Gill A. Optimisation of the technical object maintenance system taking account of risk analysis results. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (3): 420-431, https://doi.org/10.17531/ein.2017.3.13.
14. Hajkowski J, Popielarski P, Sika R. Prediction of HPDC casting properties made of AlSi9Cu3 alloy, Advances in Manufacturing, SPRINGER, Manufacturing 2017, 621-631, https://doi.org/10.1007/978-3-319-68619-6_59.
15. Hirose H. Bias Correction for the Maximum Likelihood Estimation in Two-parameter Weibull Distribution, IEEE Transactions on Dielectrics and Electrical Insulation 1999; 6: 1, https://doi.org/10.1109/94.752011.
16. https://www.reliasoft.com/Weibull [Accessed 2018].
17. Johnson R A, Miller I, Freund J E. Probability and Statistics for Engineers, eighth ed., Pearson Education Limited Co., UK, 2014.
18. Kececioglu D. Reliability & Life Testing Handbook, PrenticeHall, Inc., Englewood Cliffs, New Jersey, 1993; 1.
19. Lawless J F. Statistical Models and Methods for Lifetime Data, second ed., Wiley, 2002, https://doi.org/10.1002/9781118033005.
20. Lawless J F. Statistical Models And Methods for Lifetime Data, John Wiley & Sons, Inc., New York, 1982.
21. Lee E T, Wang J W. Statistical Methods for Survival. Data Analysis, John Wiley & Sons Inc; (3rd Edition), 2003, https://doi.org/10.1002/0471458546.
22. Legát V, Mošna F, Aleš Z, Jurča V. Preventive maintenance models - higher operational reliability. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (1): 134-141, https://doi.org/10.17531/ein.2017.1.19.
23. Liu J, Song B, and Zhang Y. Competing failure model for mechanical system with multiple functional failures. Advances in Mechanical Engineering 2018, 10(5) 1-16, https://doi.org/10.1177/1687814018773155.
24. Loska A. Exploitation assessment of selected technical objects using taxonomic methods, Eksploatacja i Niezawodnosc − Maintenance and Reliability 2013; 15, 1.
25. Młynarski S, Pilch R, Smolnik M, Szybka J. Methodology of network systems reliability assessment on the example of urban transport. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (2): 278-283, https://doi.org/10.17531/ein.2018.2.14.
26. Młyńczak M. Analiza danych eksploatacyjnych w badaniach niezawodności obiektów technicznych, Zeszyty Naukowe WSOWL, 2001; 1 (159).
27. Nelson W. Applied Life Data Analysis, John Wiley & Sons, Inc., New York, 1982, https://doi.org/10.1002/0471725234.
28. Perz P, Malujda I, Wilczyński D, Tarkowski P. Methods of controlling a hybrid positioning system using LabVIEW, 21th Scientific Polish-Slovak Conference "Machine Modeling and Simulations 2016", Procedia Engineering 2017; 177, 339-346, https://doi.org/10.1016/j.proeng.2017.02.235.
29. Pieniak D, Niewczas A M, Niewczas A, Bieniaś J. Analysis of Survival Probability and Reliability of the Tooth-composite Filling System. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2011; 2(50): 25-34.
30. ReliaSoft Corporation, Life Data (Weibull) Analysis Reference, ReliaSoft Publishing Tucson, AZ, 2008.
31. Research Project "Increase in the efficiency of functioning of public means of transport as a result of implementation of LCC and RAMS concepts in accordance with the IRIS standards based on integrated information technology system" financed by Polish National Center for Research and Development. No. PBS3/B6/30/2015.
32. Rojek I, Kujawińska A, Hamrol A, Rogalewicz M. Artificial neural networks as a means for making process control charts user friendly. In: Burduk A., Mazurkiewicz D. (eds.), Intelligent Systems in Production Engineering and Maintenance - ISPEM 2017, Advances in Intelligent

Systems and Computing, Springer, 637, 168-178, 2017, https://doi.org/10.1007/978-3-319-64465-3_17.
33. Selech J. Prognozowanie kosztów obsługiwania korekcyjnego pojazdów transportu masowego [Forecasting costs of corrective maintenance of mass transport vehicles]. Wydawnictwo Naukowe ITeE-PIB, Radom 2019, ISBN 978-83-7789-557-3.
34. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks, Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (2), 292-299, https://doi.org/10.17531/ein.2018.2.16.
35. Trojanowska J, Kolinski A, Galusik D, Varela M L R, Machado J. A methodology of improvement of manufacturing productivity through increasing operational efficiency of the production process. In: Hamrol A., Ciszak O., Legutko S., Jurczyk M. (eds) Advances in Manufacturing. Lecture Notes in Mechanical Engineering. Springer, Cham, 2018; 23-32, https://doi.org/10.1007/978-3-319-68619-6_3.
36. Waluś K J. Driver's Strategy and Braking Distance in Winter, Transport Means 2017: Proceedings of the 21st International Scientific Conference, Juodkrante, Lithuania. 2017; Part 2, 505 - 509, ISSN 1822-296 X, e-ISSN 2351-7034.
37. Wojtkowiak D, Talaśka K, Malujda I, Domek G. Estimation of the perforation force for polymer composite conveyor belts taking into consideration the shape of the piercing punch. The International Journal of Advanced Manufacturing Technology 2018, https://doi.org/10.1007/s00170-018-2381-3.
38. Ziółkowski J, Borucka A, Model Markowa w logistycznym zarządzaniu przedsiębiorstwem [Markov model in logistic management of enterprise], Journal of Konbin 2016; 2 (38), https://doi.org/10.1515/jok-2016-0027.
39. Żurek J, Ziółkowski J, Borucka A. Application of Markov processes to the method for analysis of combat vehicle operation in the aspect of their availability and readiness, Safety and Reliability - Theory and Applications - Čepin & Briš (Eds)©, Taylor & Francis Group, London, 2017; 2343-2352.

Jarosław SELECH
Karol ANDRZEJCZAK
Poznan University of Technology,
ul. Piotrowo 3A, 60-965 Poznań, Poland

E-mails: jaroslaw.selech@put.poznan.pl,
karol.andrzejczak@put.poznan.pl