

Statistical analysis of the vehicles main dimensions variability on public roads

Igor Maystrenko¹[0000-0003-0281-3112], Tagir Zinnurov^{1*}[0000-0002-7238-2883], Tatyana Maystrenko¹, and Dmiy Erochin¹

¹Kazan State University of Architecture and Engineering, 420043, Zelenaya st., Kazan, Russia

Abstract. The aim of the study is to obtain the parameters of vehicle size variability and evaluate the convergence of empirical data with the modeling results using distribution functions. The main results of the study consist in obtaining statistical parameters that characterize the variability of vehicle sizes, and testing hypotheses about whether the empirical distribution belongs to one of the theoretical distribution functions. The significance of the obtained results lies in the possibility of using the most suitable theoretical functions of probability distributions of random variables that characterize the variability of vehicle sizes in algorithms for estimating the resource parameters of steel spans of road bridges.

Keywords: statistical, bridge, distribution functions, vehicle sizes, model.

1 Introduction

The practical state of the traffic flows shows statistical studies issue that the main dimensions of vehicles (length, wheelbase, track and width) have a certain statistical variability.

In the work of Drew D., it was proposed to use the parameters of four types of calculated cars – passenger, cargo, tractor-trailer with a medium semi-trailer, tractor-trailer with a large semi-trailer-the total length of which is 5.79 m., 9.14 m., 15.24 m., 16.76 m, respectively [1].

Abdunazarov proposed the parameters of the recommended design cars for the design of highways. However, his works do not provide a justification for the accepted law of distribution of a random variable for describing these parameters [2, 3]. It is proposed to take the value of 4.9 m as the calculated value of the length for a passenger car, and 6.8 m for a truck. This means that the proposed calculated values of the size of vehicles do not allow the design of structures on highways with a given level of confidence.

The variability of the cars and other vehicles parameters are considered at the level of national standards and reflected in the reference literature [4-6]. Although the systematized data are proposed as part of the assessment of the safety of vehicle operation, the information obtained can be used in the formation of mathematical models at various levels.

*Corresponding author: leongar@mail.ru

The research is conducted in a direction related to the selection and justification of the distribution laws for further modeling. The authors Haq S., Temkin M., Black L. and Bammel P. present methods that are used in the development of road simulation models to determine the fluctuations in road load [7]. The method of statistical analysis of the most representative load profile from a set of time cycles is considered.

The researcher Yakubovich A. N. proposed an algorithm for constructing a sample from realizations of a random variable for which its probability density is known [8]. The proposed algorithm does not use pseudorandom number generators and is characterized by high accuracy to reproduce the simulated distribution law with a small number of implementations.

Mazlounian A., Geroliminis N., and Helbing D. reveal fundamental relationships between mean flow, mean density, and vehicle density variability [9]. The motion inhomogeneities are considered as an independent variable, which makes it possible to exclude the scattering of an overloaded flow measurements.

The authors Jiwon K., Hani S. M. propose a comprehensive probability distribution approach to account for the inter-vehicle and daily variability of traffic flow in the road network [10]. When modeling the traffic delay, a gamma model is proposed. The main advantage of the model is its ability to recognize different sizes of flow variability, which are reflected in the time cycles of movement.

The study of the variability of physical parameters using various laws is described in the works of the authors Chen P., Ansari Esfeh M., Chiou J. M., Manapov A. Z. and others [11-16]. There are other works in this area of research [17-21].

It is obvious that for passenger cars there is a steady trend towards a certain increase in the size of vehicles with the release of updated models. For trucks, the variability in size is often associated with an increase in traffic volumes and with intensive cooperation at the international level, when automobile factories of foreign manufacturers are opened in different regions.

In this paper, the authors conducted a statistical study aimed at finding acceptable theoretical functions of probability distributions that allow us to take into account the variability of the main dimensions of vehicles on public roads when conducting a multi-factor simulation of the operation of steel spans of road bridges.

2 Materials and methods

In order to obtain the parameters of vehicle size variability, the authors conducted a statistical analysis, during which samples of the most common vehicle models were compiled, grouped by type: the first group – passenger cars; the second group-trucks, including single cars, semi-trailers and automobile trains.

Considering that at the moment there is no single international classification of vehicles. The statistical analysis of the passenger cars size variability is based on the accepted segmentation in the Economic Commission for Europe. The most common passenger car brands were grouped into the following segments: A – «mini cars», B – «small cars», C – «medium cars», D – «large cars», E – «executive cars», F – «luxury cars», S – «sport coupes», M – «multi-purpose cars», J – «sport utility cars (including off-road vehicles) ». The trucks were grouped into four segments: MT – «low-tonnage single trucks, carrying capacity up to 3 tons»; CT – «medium-tonnage single trucks, carrying capacity from 3 to 5 tons, body volume not less than 35 m³»; CT-1 – «large-tonnage trucks, including semi-trailers and trailed road trains, carrying capacity from 5 tons, body volume not less than 80 m³, weighing up to 20 tons inclusive»; KT-2 – «large-capacity trucks, including semi-trailers and trailed road trains, carrying capacity from 5 tons, body volume of at least 80 m³, weight from 20 to 44 tons inclusive». The total volume of samples was 211 passenger car

brands and 261 modifications of various manufacturers trucks. The statistical parameters that characterize the variability in the size of vehicles are shown in table 1.

Table 1. Statistical parameters that characterize the variability of vehicle sizes.

Vehicle parameters	Cross-group and generalized statistical characteristics for the entire sample data set					
	$\mu_{(*)T}$, m	$D_{(*)Tzs}$, m ²	$D_{(*)Tz}$, m ²	$D_{(*)T}$, m ²	$\Delta_{(*)T}$, m	$\nu_{(*)T}$
Segments A, B, C, D, E, F, S, M, J						
Passenger car length $L_{T,1}$	4.542	0.081	0.254	0.335	0.579	0.127
Passenger car wheelbase $d_{T,1}$	2.726	0.023	0.058	0.081	0.285	0.104
Passenger car tread $K_{T,1}$	1.552	0.003	0.006	0.009	0.093	0.060
Passenger car width $B_{T,1}$	1.813	0.005	0.009	0.014	0.116	0.064
Segments MT, CT, KT-1, KT-2						
Trucklength $L_{T,2}$	8.857	3.300	51.783	55.083	7.422	0.838
Truck wheelbase $d_{T,2}$	4.993	1.860	14.760	16.620	4.077	0.816
Trucktread $K_{T,2}$	1.842	0.019	2.904	2.923	1.710	0.928
Truck width $B_{T,2}$	2.276	0.020	4.554	4.574	2.139	0.940

The following designations are accepted: $\mu_{(*)T}$ – general mean for the whole population; $D_{(*)Tzs}$ – intra–group variance; $D_{(*)Tz}$ – inter–group variance; $D_{(*)T}$ – generalized variance; $\Delta_{(*)T}$ – standard deviation for the whole population; $\nu_{(*)T}$ – coefficient of variation for the whole population [22].

To describe the parameters that characterize the variability of vehicle dimensions $(*)T$, four competing probability distribution functions were considered, with their characteristic statistical estimates, the values of which are presented in table 2.

Table 2. Random variable distribution functions.

Distribution	Probability density function of the distribution
Normal	$F_N[(*)T] = \frac{1}{\Delta_{(*)T} \sqrt{2\pi}} \exp \left\{ -\frac{[(*)T - \hat{\mu}_{(*)T}]^2}{2\Delta_{(*)T}^2} \right\}$
Extreme Gumbel values (Type I)	$F_G[(*)T] = \frac{1}{\lambda_{(*)T}} \exp \left[\frac{\mu_{(*)T} - (*)T}{\lambda_{(*)T}} - \exp \left(\frac{\mu_{(*)T} - (*)T}{\lambda_{(*)T}} \right) \right]^2$
Rayleigh	$F_R[(*)T] = \frac{(*)T}{\beta_{(*)T}^2} \exp \left(-\frac{(*)T^2}{2\beta_{(*)T}^2} \right)$
Uniform	$F_U[(*)T] = \begin{cases} \frac{1}{b_{(*)T} - a_{(*)T}}, & (*)T \in [a_{(*)T}, b_{(*)T}] \\ 0, & (*)T \notin [a_{(*)T}, b_{(*)T}] \end{cases}$

The results of estimating the convergence of theoretical and empirical distributions of random variables of vehicle size variability are performed using a sample correlation coefficient $r_{(*)T}$ with the construction of a confidence interval $CI_{(*)T}$ (Table 3).

Table 3. Statistical characteristics of generalized data samples (for all segments and groups) that characterize the variability of vehicle sizes.

Parameter	n_{Σ}	$\mu_{(*)T}$, m	$\Delta_{(*)T}$, m	$\nu_{(*)T}$	$med_{(*)T}$	$mod_{(*)T}$
Vehicle length L_T	472	6.928	4.433	0.640	5.801	19.825
Vehicle wheelbase d_T		4.329	3.275	0.757	3.350	3.800
Vehicle tread K_T		1.748	0.323	0.185	1.683	2.032
Vehicle width B_T		2.111	0.445	0.211	2.010	2.500
<i>med</i> (*)T - sample median; <i>mod</i> (*)T - sample moda						

To construct the confidence interval $CI_{(*)T}$ for the correlation coefficient $r_{(*)T}$, obtained from a small sample divided into Requal-length segments, the well-known Fisher transform is used:

$$z = arth r_{(*)T} = 0.5 \cdot \ln \frac{1+r_{(*)T}}{1-r_{(*)T}} \tag{1}$$

where $arth r_{(*)T}$ – inverse hyperbolic tangent function for the correlation coefficient $r_{(*)T}$.

Let's check the null hypothesis $H_0[(*T)]$ that the empirical distribution obeys one of the theoretical ones constructed from a sample of volume n_{Σ} . The test was performed using the Pearson's agreement criterion χ^2 for the significance level $\alpha = 0.05$ [22]

$$\chi^2 = \frac{\sum (n_i - n'_i)^2}{n'_i} \tag{2}$$

Where $n_i = n_1, n_2, \dots, n_R, n'_i = n'_1, n'_2, \dots, n'_R$ - empirical and theoretical frequencies, respectively.

The results of testing the hypothesis that the sample belongs to a certain distribution law are presented in table 4.

Table 4. Statistical estimates of the parameters of the distribution functions used to describe the variability of vehicle dimensions.

Parameter	Statistical estimates for accepted distribution functions			
	Normal distribution $F_N [(*T)]$	Rayleigh distribution $F_R [(*T)]$	Distribution Gumbel $F_G [(*T)]$	Uniform distribution $F_U [(*T)]$
L_T , m	$\hat{\mu}_{LT} = 6.928$ $\Delta_{LT} = 4.433$	$\beta_{LT} = 5.528$	$\mu_{LT} = 4.932$ $\lambda_{LT} = 3.456$	$\alpha_{LT} = 2.285$ $\beta_{LT} = 19.905$
d_T , m	$\hat{\mu}_{dT} = 4.329$ $\Delta_{dT} = 3.275$	$\beta_{dT} = 3.454$	$\mu_{dT} = 2.854$ $\lambda_{dT} = 2.553$	$\alpha_{dT} = 1.500$ $\beta_{dT} = 18.900$
K_T , m	$\hat{\mu}_{KT} = 1.748$ $\Delta_{KT} = 0.323$	$\beta_{KT} = 1.395$	$\mu_{KT} = 1.602$ $\lambda_{KT} = 0.252$	$\alpha_{KT} = 1.200$ $\beta_{KT} = 2.330$
B_T , m	$\hat{\mu}_{BT} = 2.111$ $\Delta_{BT} = 0.445$	$\beta_{BT} = 1.684$	$\mu_{BT} = 1.910$ $\lambda_{BT} = 0.347$	$\alpha_{BT} = 1.380$ $\beta_{BT} = 2.722$

Table 5. Results of the estimation of the convergence of theoretical and empirical distributions of the random variable of vehicle size variability by the sample Pearson correlation coefficient for a confidence probability of 0.95.

Parameter	The value of the Pearson sample correlation coefficient $r_{(*)T}$ / confidence interval $CI_{(*)T}$ for the correlation coefficient for the corresponding theoretical distribution functions			
	Normal distribution $F_N[(*)T]$	Rayleigh distribution $F_R[(*)T]$	Distribution Gumbel $F_G[(*)T]$	Uniform distribution $F_U[(*)T]$
L_T, m	0.534/ (0.119; 0.703)	0.660/ (0.207; 0.746)	0.695/ (0.230; 0.757)	-0.004/ (0.364; 0.357)
d_T, m	0.554/ (0.134; 0.711)	0.687/ (0.225; 0.755)	0.692/ (0.228; 0.756)	0.081/ (-0.287; 0.430)
K_T, m	0.665/ (0.211; 0.748)	0.457/ (0.057; 0.670)	0.735/ (0.254; 0.767)	-0.035/ (-0.391; 0.3296)
B_T, m	0.440/ (0.043; 0.663)	0.221/ (-0.155; 0.536)	0.467/ (0.066; 0.675)	-0.116/ (-0.458; 0.254)

Table 6. Results of testing hypotheses about whether an empirical distribution belongs to one of the theoretical distribution functions according to the Pearson agreement criterion χ^2 .

Parameter	Theoretical probability distribution functions used to describe parameters that characterize the variability of vehicle dimensions			
	Normal distribution $F_N[(*)T]$	Rayleigh distribution $F_R[(*)T]$	Distribution Gumbel $F_G[(*)T]$	Uniform distribution $F_U[(*)T]$
L_T, m	2.672	2.133	1.489	6.265
d_T, m	4.290	16.332	1.893	7.418
K_T, m	12.833	45.992	12.514	955.122
B_T, m	26.223	67.240	30.897	1273.283

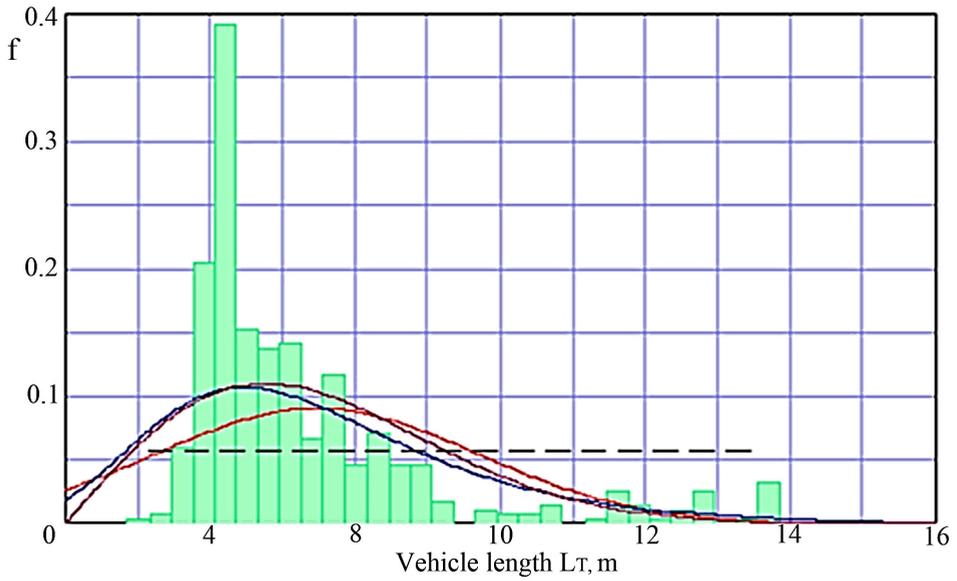


Fig. 1. Frequency diagram of statistical data φ and theoretical distribution density functions f of the parameter L_T .

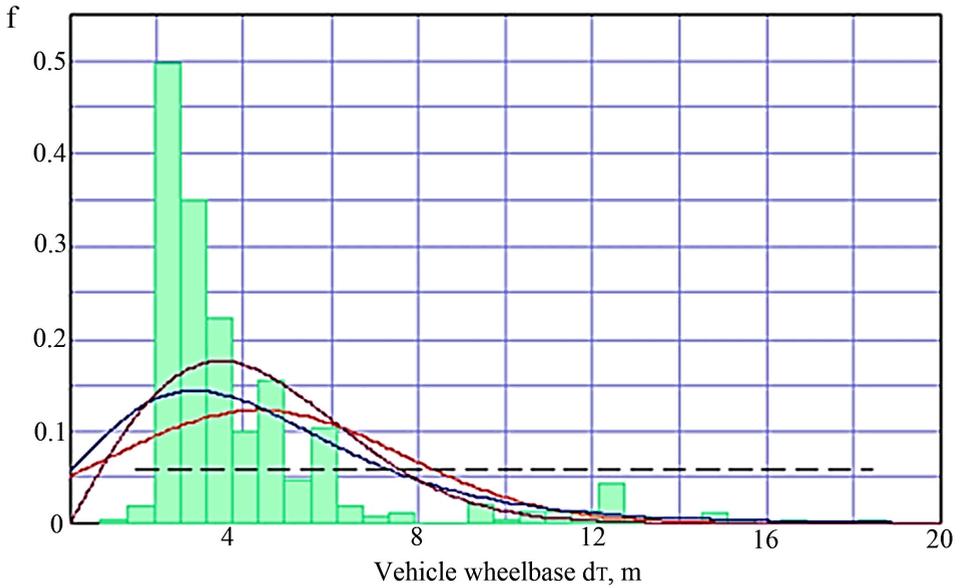


Fig. 2. Frequency diagram of statistical data φ and theoretical distribution density functions f of the parameter d_T .

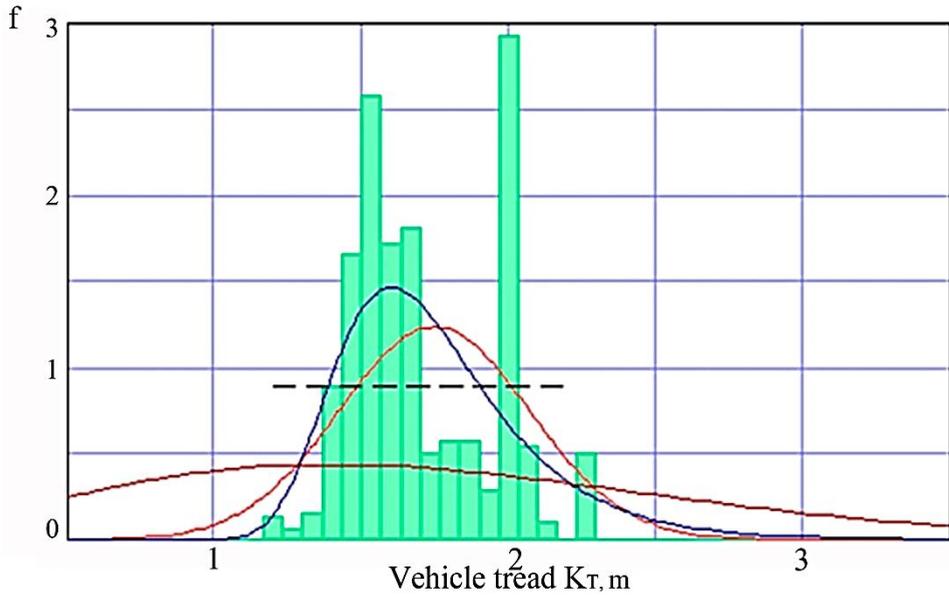


Fig. 3. Frequency diagram of statistical data φ and theoretical distribution density functions f of the parameter K_T .

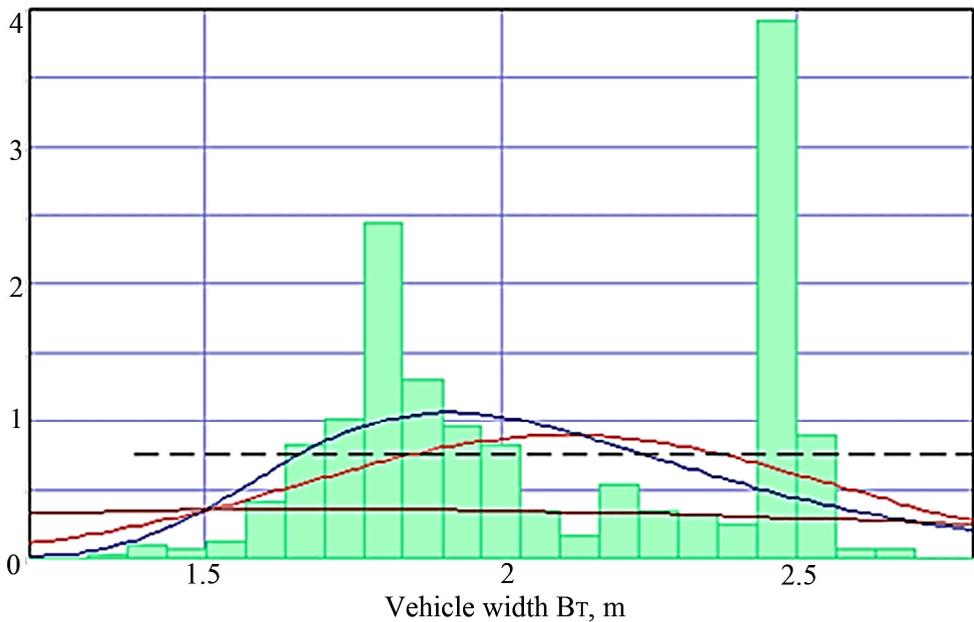


Fig. 4. Frequency diagram of statistical data φ and theoretical distribution density functions f of the parameter B_T .

In the graphs of figures 1-4, the red solid line is the density of the distribution according to the normal law, the brown solid line is the density of the distribution according to the Rayleigh law; the blue solid line is the density of the distribution according to the Gumbel

law (Type I), the black dotted line is the density of the distribution according to the uniform law, the green diagram is statistical data.

3 Results and discussion

The results of the analysis of the data of the statistical study conducted to evaluate the parameters of the distribution functions that can be used to describe the variability of vehicle sizes, presented in figures 1-4 and in tables 3-6, showed:

Parameter «car length», L_T . To describe the L_T parameter, the theoretical Gumbel distribution function with the position parameter (mode) $\mu_{L_T} = 4.932$ m and the scale parameter $\lambda_{L_T} = 3.456$ m is adopted, taking into account the existing restrictions on the maximum length of road trains of 20 meters for public roads. For the accepted distribution function, the sample Pearson correlation coefficient has the highest value compared to other alternative theoretical probability distribution functions considered and is within the range of $0.5 < r_{L_T} = 0.695 < 0.7$, while the confidence interval with a reliability of 0.95 for the correlation coefficient $CI_{L_T} = (0.230; 0.757)$ does not pass through zero – therefore, we can judge the presence of a direct noticeable linear correlated relationship on the Cheddock scale between the empirical and accepted theoretical distribution functions of the calculated parameter L_T .

The «car wheelbase» parameter, d_T . To describe the d_T parameter, the theoretical Gumbel distribution function with the position parameter (mode) $\mu_{d_T} = 2.854$ m. and the scale parameter $\lambda_{d_T} = 2.553$ m. For the accepted distribution function, the Pearson sample correlation coefficient has the highest value compared to other alternative theoretical probability distribution functions (competing hypotheses) considered and is within the range of $0.5 < r_{d_T} = 0.692 < 0.7$, while the confidence interval with a reliability of 0.95 for the correlation coefficient $CI_{d_T} = (0.228; 0.756)$ does not pass through zero – therefore, we can judge the presence of a direct noticeable linear correlated relationship on the Cheddock scale between the empirical and accepted theoretical distribution functions of the calculated parameter d_T .

Parameter «car track», K_T . To describe the K_T parameter, the theoretical Gumbel distribution function with the position parameter (mode) $\mu_{K_T} = 2.854$ m and the scale parameter $\lambda_{K_T} = 2.553$ m is adopted. For the accepted distribution function, the Pearson sample correlation coefficient has the highest value compared to other alternative theoretical probability distribution functions considered (competing hypotheses) and is within the range of $0.7 < r_{K_T} = 0.735 < 0.9$, while the confidence interval with a reliability of 0.95 for the correlation coefficient $CI_{K_T} = (0.254; 0.767)$ does not pass through zero – therefore, we can judge the presence of a direct high linear correlated relationship on the Cheddock scale between the empirical and accepted theoretical distribution functions of the calculated parameter K_T .

The «car width» parameter, B_T . To describe the B_T parameter. the theoretical function of the Normal distribution is adopted, with the mathematical expectation parameter $\mu_{B_T} = 2.111$ m. and the variance $\Delta_{B_T} = 0.445$, taking into account the existing restrictions on the largest width of cars 2.55 m. for public roads. For the accepted distribution function, the Pearson sample correlation coefficient has the highest value compared to other alternative theoretical probability distribution functions considered (competing hypotheses) and is in the range of $0.3 < r_{B_T} = 0.440 < 0.5$, while the confidence interval with a reliability of 0.95 for the correlation coefficient $CI_{B_T} = (0.043; 0.663)$ does not pass through zero – therefore, we can judge the presence of a direct moderate linear correlated relationship on the Cheddock scale between the empirical and accepted theoretical distribution functions of the calculated parameter B_T .

4 Conclusions

The statistical analysis of the variability of the main dimensions of vehicles (length, wheelbase, track and width) showed that the following functions can be used to develop an algorithm for calculating the steel spans of road bridges based on the probability of failure-free operation in the first approximation:

- the distribution function of the extreme value (type I Gumbel) with the following parameters provisions $\mu_{(*)T}$ and scale $\lambda_{(*)T}$: to describe the size of the «length of the car» – $\mu_{LT} = 4.932$ m and $\lambda_{LT} = 3.456$ m; to describe size «wheelbase of the car» – $\mu_{dT} = 2.854$ m and $\lambda_{dT} = 2.553$ m.; to describe the size of a «track car» – $\mu_{KT} = 2.854$ m and $\lambda_{KT} = 2.553$ m.;
- function of the normal distribution with the following parameters $\mu_{(*)T}$ and $D_{(*)T}$ – scores of mathematical expectations and variances and variances to describe the size of the «width of the car» – $\mu_{BT} = 2.111$ m., $\Delta_{BT} = 0.445$.

When using the asymptotic distribution to create arrays of calculated parameters pseudo-random implementations that characterize the main dimensions of vehicles, it is necessary to introduce boundary conditions that take into account the physical nature of the calculated factors, as well as the requirements of the design standards for public roads. The noted imperfection of the asymptotic distribution can be eliminated by converting it to a truncated distribution. The transformation is performed by reallocating events from the region of impossible values for a particular physical parameter to the region of possible values.

References

1. D. Drew, *Traffic flow theory and management*, Transport, **424** (1972)
2. Zh. N. Abdunazarov, Substantiation of the parameters of calculated vehicles in the design of geometric elements of highways (Abstract dis. for the degree of candidate of technical sciences. (05.23.11) MAD120, 2015)
3. Zh. N.nAbdunazarov, M. Mamarasulova, *Recommended parameters of design vehicles for the Russian Federation*, International scientific journal «Molodoyiucheniy», **7.2 (111.2)** (2016)
4. Improving the Compatibility of Vehicles and Roadside Safety Hardware National Academies of Sciences, Engineering, and Medicine 2004. Improving the Compatibility of Vehicles and Roadside Safety Hardware. Washington, DC: The National Academies Press(2004)
5. Surface vehicle recommended practice, SAEJ1100 rev. (2001)
6. R. S. Foot, *Installation of a Tunnel Traffic Surveillance and Control System*, Traffic Control Theory and Instrumentation, (1965)
7. S. Haq, M. Temkin, L. Black, P. Bammel, *Vehicle road simulation testing, correlation and variability*, SAE Technical Paper, 2005-01-0856 (2005) DOI:10.4271/2005-01-0856
8. A. N. Yakubovich, I. A. Yakubovich, *An effective algorithm for the numerical implementation of random variables in the problems of the technical systems reliability statistical modeling*, Intelkt. Innovatsii. Investitsii **8** (2016)
9. A. Mazlounian, N. Geroliminis, D. Helbing, *The spatial variability of vehicle densities as determinant of urban network capacity*, Philos Trans A Math Phys Eng Sci. (2010)
10. J. Kim, H. S. Mahmassani, *Compound gamma representation for modeling travel time variability in a traffic network*, Transportation Research Part B: Methodological, **80** (2015) DOI: 10.1098/rsta.2010.0099
11. P. Chen, R. Tong, G. Lu, Y. Wang, *Exploring travel time distribution and variability patterns using probe vehicle data: Case study in Beijing*, Journal of Advanced Transportation (2018) DOI: 10.1155/2018/3747632

12. J. M. Chiou, H. T. Liou, W. H. Chen, *Modeling time-varying variability and reliability of freeway travel time using functional principal component analysis*, IEEE Transactions on Intelligent Transportation Systems, **22** (2021) DOI:10.1109/TITS.2019.2956090
13. M. Ansari Esfeh, L. Kattan, W. Lam, R. Ansari Esfe, M. Salari, *Compound generalized extreme value distribution for modeling the effects of monthly and seasonal variation on the extreme travel delays for vulnerability analysis of road network*, Transportation Research Part C:Emerging Technologies (2020) DOI:10.1016/S0165-1765(03)00035-1
14. A. Z. Manapov, *Truncated normal distributions and their application in solving problems of structural reliability*, IzvestiaVUZov. Stoitelstvo i architectura, **9**(1988)
15. A. Z. Manapov, *Calculation of structural reliability using truncated distributions of strength and stress*, IzvestiaKGASU, **2(6)** (2006)
16. E. V. Chimitova, E. O. Ermilova, *Investigation of the properties of estimates of the maximum likelihood of the parameters of the Weibull distribution based on data truncated from the left*, VestnikSibGUTI, **4** (2016)
17. R. L. Sakhapov, R. V. Nikolaeva, M. H. Gatiyatullin, M. M. Makhmutov, *Risk management model in road transport systems*, Journal of Physics: Conference Series, **738(1)** (2016) DOI: 10.1088/1742-6596/738/1/012008
18. I. Yu. Maystrenko, T. A. Zinnurov, D. I. Erohin, T. I. Maystrenko, *Analysis of the efficiency of obtaining reliable data when operating with number sets in simulation of bridges structural components*, IOP conference series: Materials Science and Engineering, **890**, 012030 (2020) DOI: 10.1088/1757-899X/890/1/012030
19. E. Vdovin, L. Mavliev, V. Stroganov, *IOP Conf. Ser. Mater. Sci. Eng.* (2020)
20. I. V. Grishyn, G. P. Ivanov, R. A. Kayumov, *Durability of bridge asphaltic concrete pavements under temperature loads*, IOP Conf. Series: Materials Science and Engineering, **786(1)**, 012032 (2020) DOI: 10.1088/1757-899X/890/1/012032
21. R. Galeev, R. Nizamov, L. Abdrakhmanova, V. Khozin, in *IOP Conf. Ser. Mater. Sci. Eng.* (2020)
22. V. D. Kelton, A. M. Loy *Imitation modeling* (Classika CS, Piter, Saint-Petersburg, Russia, 2004)