

Andrzej KORNACKI
Jacek WAWRZOSEK
Andrzej BOCHNIAK
Andrzej SZYMANEK
Halina PAWLAK

CRITICAL VALUES OF DRIVER RESPONSE TIME AND ITS IMPACT ON REDUCING RELIABILITY AND SAFETY IN ROAD TRAFFIC

KRYTYCZNE WARTOŚCI CZASU REAKCJI KIEROWCY I ICH WPŁYW NA OBNIŻENIE NIEZAWODNOŚCI I BEZPIECZEŃSTWA RUCHU DROGOWEGO*

Road traffic is among the most dangerous types of human activity. The main causes of road accidents are driver fatigue, poor physical and mental condition of drivers and overestimating one's skills while driving. This study focuses on the estimation of driver response time, as the basis of a hypothetical system that uses short and long-range radars, which determines the physical and mental condition of a driver; based on the analysis of „acceleration noise” of the vehicle following its predecessor. This work highlights serious consequences of the fact that driver response time is described by means of a distribution with heavy tails, and thus may be a source of hazard in the driver-vehicle system. Extremes of driver response time were treated as outliers in this study. Their detection was attained by using the Akaike information criterion [1, 2], which is an alternative to conventional methods of testing hypotheses. Untypical, on account of their outlying nature, values are interpreted as critical driver response time values which potentially endanger the reliability of driving.

Keywords: *driver response time, reliability of road traffic, outliers, Akaike information criterion, log-normal distribution.*

Ruch drogowy należy do najbardziej niebezpiecznych rodzajów działalności człowieka. Główne przyczyny wypadków drogowych to zmęczenie kierowców, zły stan psychofizyczny kierujących oraz przecenianie swoich umiejętności podczas prowadzenia pojazdu. W niniejszej pracy skupiono uwagę na estymacji czasu reakcji kierowców, jako podstawie hipotetycznego systemu wykorzystującego radary dalekiego i krótkiego zasięgu a określającego stan psychofizyczny kierowcy w oparciu o analizę „szumu przyspieszeń” pojazdu podążającego za poprzednikiem. Wskazuje się na groźne konsekwencje faktu, że czas reakcji kierowcy jest opisywany rozkładem z ciężkimi ogonami, gdyż z tego powodu może być źródłem zagrożenia w układzie kierowca-pojazd. Skrajne wartości czasu reakcji kierowców potraktowano w pracy, jako wartości odstające. Do ich wykrycia zastosowano kryterium informacyjne Akaike [1, 2] co stanowi alternatywę w stosunku do klasycznych metod testowania hipotez. Nietypowe, bo odstające wartości interpretuje się, jako krytyczne czasy reakcji kierowców potencjalnie zagrażające niezawodności jazdy.

Słowa kluczowe: *czas reakcji kierowców, niezawodność ruchu drogowego, obserwacje odstające, kryterium informacyjne Akaike, rozkład logarytmiczno-normalny.*

1. Introduction

Road traffic is one of the most dangerous types of human activity. It is due to several factors, including the unreliability of the driver-vehicle system in the process of driving. According to WHO statistics, 1.2 million of people die in car accidents each year. In the past 10 years on EU roads ca 0.5 million people were killed with over 1.5 million seriously injured and many people remain disabled for the rest of their lives. The annual cost of road accidents in the EU amounts to € 160 billion, which constitutes 2% of EU GDP. The risk of being a fatality on Polish domestic roads is four times greater than on roads in Germany and the UK, and on highways it is even six times greater. In Poland, in 2005, as many as 57.3% of the accidents occurred on a straight stretch of road, and the most common cause of road accidents was excessive speed (28.8%), while failure to maintain a safe distance between vehicles causes 4.5% of accidents. In 2012, about 11%

of accidents are the result of driving into the rear of another vehicle because of not maintaining a safe distance. The cause of up to one quarter of accidents on German highways is driver drowsiness. Most of these accidents have severe consequences. The main risk factors on roads are: human, as a participant of the traffic, the vehicle, and the road. As many as 90-95% of accidents are related to man and their behaviour. The vehicle contributes to causing 8-10% of road accidents, while the road and its surroundings contributes to 28-35% of road accident occurrence. These factors may impact the risk individually or several factors may combine (Fig. 1) [14, 20]. Archer [3] observed that at an average speed of 60 km/h average one risky situation for 120 km should be attributed to an average driver.

Human behaviour, ignorance of one's own body, response to stress or the level of fatigue or excessive confidence in one's own driving capability contribute to the occurrence of accidents. Other factors, such as weather conditions or the road surface condition, are much

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

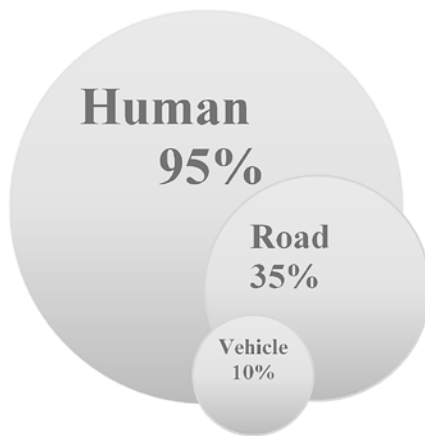


Fig. 1. The influence of the major risk factors on the incidence of road accidents.

less significant. Thus, in many centres there is ongoing development of integrated monitoring systems for the psycho-physical condition of drivers. An overview of some systems enhancing road safety, currently deployed and being in development, is included in works of Wicher [34], Mercedes-Benz [27], Cieślak and Karpińska [9]. They illustrate the extensive engineering work of, among others, SAAB, Mitsubishi, Toyota, Mercedes-Benz, Nissan, Lexus companies. Vehicle electronics is the fastest developing segment in consumer electronics. In 1980, the value of electronic vehicle equipment accounted for less than 1% of the vehicle price. In 1990, it was already close to 7%, and in 2007, this share increased to 22%. Current estimates exceed 40% of the vehicle value. The value of electronic vehicle control systems is currently growing at a rate of ten plus percent per annum.

Theoretical considerations presented in this article highlight the need and create the possibility to further develop the concept of yet another system, which this time monitors components of the psycho-physiological response time of the driver with regard to calculating accident risk in the process of dependent driving, which is the primary component of traffic flows.

2. System-based, on-going control of driver response time

Individual physiological characteristics define the limitations of the driver's organism. Engineering psychology involves checking the following psycho-physical parameters of drivers: response time and its consistency, assessing vehicle speed, spatial vision, sensitivity of proprioception, sensitivity to glare and the ability to distinguish shapes in the darkness, speed and accuracy of perception, speed and accuracy of decision making, divisibility and concentration of attention, technical knowledge, resistance to fatigue, neuroticism [23]. The following are among the most important characteristics of a driver: quality of vision, response time and age.

Response time constitutes part of the time needed to stop a car. The manoeuvre of stopping a car before an appearing obstacle can be divided into the following: perception time, actual driver response time, time required to activate the braking system of the car and the actual braking time. In some publications perception time, which is the time that elapses between the moment an obstacle can be noticed and the moment the central nervous system begins working, is included in the time slot called response time. Similarly, to this study, it stems from the conditions under which an experiment is conducted. The combined driver's perception and response times may change, ranging from 1 to even a few seconds depending on:

- driver's current health condition, mood, and physical characteristics (e.g. headache, being well-rested or tired),

- driver's degree of concentration (long and monotonous drive or prolonged noise in the driver's compartment can increase the response time by more than 10%),
- driver's circadian rhythm (decreased vigilance, fatigue, drowsiness), time of year, day and meteorological phenomena (changes in atmospheric pressure and temperature of the environment, precipitation),
- ergonomic conditions (e.g. inappropriate placement of the control system elements, poorly adjusted parameters of the auxiliary subsystems),
- complexity of the traffic situation (surprise as a result of an unusual road situation), the number of options considered by the driver and the actual manoeuvres undertaken,
- street lighting, type of road (street, suburban road, motorway),
- insobriety, influence of psychotropics, intoxicating substances.

If the driver anticipates danger, it is assumed that the response time is 0.5-0.8 s. If the driver is careful but does not anticipate danger, it is 0.7-0.9 s. When the driver steers carelessly, it is 1.4-1.9 s. Computer reconstructions of road collisions often assume that the total perception and driver response time is 1.75 s. During night time this duration is longer and amounts to 2.5 s [5].

Response times estimated on the basis of laboratory test simulations ought to be taken very carefully. Most of these tests consist in measuring driver response time to a simple signal, i.e. such as measuring the time that elapses between the appearance of red light and pressing the brake pedal. These tests do not reflect the real conditions encountered by a driver in traffic and the measured response times are much shorter [35].

The constituent elements of the currently installed optional equipment of the vehicle, i.e. a combination of a stereoscopic camera with long and short range radar sensors, after appropriate reprogramming, allow for direct measurements of the response time of a driver following another vehicle. Multiple measurements of the time elapsing between the moment the preceding vehicle brakes and the moment of the following vehicle driver's response can be a parameter that may be used in a variety of ways. It can, for example, be used to warn the driver about their poor psycho-physical condition, of the need to change their driving style or the parameters of auxiliary subsystems, it can also lead to reducing the speed of the vehicle, temporary exclusion from the traffic, signalling a problem to other users and even notifying the police about a threat.

3. Critical driver response time

In order to compare different traffic situations, Jones and Potts [21] proposed a probabilistic description of traffic flow, where a significant role was assigned to the parameter of σ_a – acceleration dispersion, commonly referred to as “acceleration noise” of a vehicle following its predecessor (the so-called dependent driving). Model testing has utilised this traffic quality indicator. Acceleration noise incorporates both the longitudinal and the transverse component. The latter is particularly visible on a winding road, but it has not yet been adequately examined [22]. Usually only the longitudinal component expressed by the following parameter is considered: σ_a – acceleration dispersion [4, 35]. The σ_a indicator was verified in simulation tests, and the results confirmed its correlation with statistical safety indicators. Effective research on real objects has not been yet conducted. The tragic death of the Polish precursor in real time traffic flow, Mieczysław Kaczor, D.Eng., discontinued the already initiated research [23].

Acceleration dispersion is a consequence of the time of the driver's response to the observed stimuli as well as their intensity, which are associated with the psycho-physical predispositions of the driver,

manifested in their driving style. Acceleration dispersion is also a performance function of the vehicles participating in road traffic.

“Acceleration noise” (sometimes referred to as “noise of accelerations”) is the mean square deviation of the acceleration distribution of the vehicles participating in road traffic [13]. This indicator defines the inability to maintain a constant drive speed. If to determine the level of “acceleration noise” a measurement method called the “floating” vehicle is used, the “acceleration noise” of a traffic flow with a small error resulting from the limited time of T measurements can be determined by measuring changes in the tested vehicle acceleration. Then:

$$\sigma_a = \lim_{T \rightarrow \infty} \sqrt{\frac{1}{T} \int_0^T [a(t) - \bar{a}]^2 dt},$$

where $a(t)$ \bar{a} – are respectively: realisations of the stochastic process $X(t)$ (temporally random changes in the tested vehicle acceleration) and the average acceleration of the tested vehicle over the measuring distance $\langle 0; T \rangle$.

The examination of the critical response time of drivers should allow us to show the relationship between potentially dangerous road traffic situations and the following categories of driver response time: typical and untypical, permissible and impermissible, safe and endangering safety. The assignment of any of these categories may refer to a single measurement, as well as a group of measurements. In the latter case, we will refer to the current characteristics of the physiological features of the driver, e.g. fatigue, intoxication. When the exemplary, typical individual characteristics of driver response time are known, they can be easily compared with characteristics of the current data stream, and then prevent a potential threat. Just like in systems monitoring driver fatigue by analysing the movements and fixation of the eyeball; they also deploy the comparison of the digital “rested eye” pattern with actual camera observations of the driver’s eye throughout the whole journey.

Each of these characteristics can be based on 1) the measurement of the percentage of stress exceeding a certain barrier or 2) filtering the data stream so that the highlighted group of multiple measurements will correspond to the actual probability of the response time distribution. By combining both methods, one can specify 3) the frequency parameter of the untypical driver response.

The system of managing the data stream coming from various sensors requires complex filtering and alarm activation in reaction to untypical observations [18]. In this study, the extreme values of driver reaction time are considered as the so-called “outlier measurements” values. While analysing the classical statistics of experimental results, it is important to determine whether the outliers are derived from a different population than the rest of the results. Then, in further analysis, untypical observations used to be customarily rejected. It is also possible, although unlikely to occur, that “strange” observations will emerge with the same distribution as for the remaining results. Then, it would be advisable to keep such observations for further statistical analysis, simultaneously increasing its effectiveness.

Methods of hypothesis testing are the most commonly used ones for the detection of outliers [6, 8, 15, 19, 30, 32]. However, conclusions in the hypothesis testing method are dependent on the assumed level of significance and may be different for its various values. Furthermore, there may appear the effect of outlier “masking”. For the data on the strength of plastic materials, Grubbs [19] describes a situation where tests do not detect one smallest observation, while two smallest observations are readily identified as outliers (a certain contradiction).

In this study, it is suggested to use the Akaike criterion information to detect outliers. This criterion, derived from information theory, allows one to select it from among the models describing experimen-

tal data, which maximizes entropy [1, 2]. According to Sakamoto [31], the value of this criterion is equal:

$$AIC = -2 \ln(\max \text{likelihood}) + 2K, \quad (1)$$

where $\max \text{likelihood}$ means likelihood calculated for the estimators of parameters obtained through the maximum likelihood method, and K the number of these parameters. We select the model for which the value of AIC is the lowest. This procedure is independent of the level of significance, the number of outliers and of whether the “suspicious” observations belong to the group of the smallest or largest observations.

4. Log-normal distribution

It appears that the log-normal distribution is well-suited for describing specific random positive values, which here take the form of the distribution of driver response time.

Definition: Random variable X has a log-normal distribution with parameters μ, σ^2 denoted $LN(\mu, \sigma^2)$ when its logarithm has normal distribution, i.e. $Y = \ln X \sim N(\mu, \sigma^2)$. Therefore, we have:

$$X \sim LN(\mu, \sigma^2) \Leftrightarrow Y = \ln X \sim N(\mu, \sigma^2). \quad (2)$$

Most probabilistic models have not considered the fact of the commonly occurring, so-called fat tails and asymmetry in the distributions of random variables describing extreme values [17 §3.2.1; 10; 16 §2]. It is estimated that the risk management methods based on the theory of extreme values do not have these disadvantages and allow one to effectively model rare but dangerous events. The log-normal distribution is often proposed for constructing models with extreme values for actual data [7 §3.3, §4.5.2]. Engineering applications of this distribution result from the fact that the description of measurement errors often must assume that the difference in the measurement result of the actual parameter value is a positive value. Examples include phenomena in which the nominal parameter values are equal to zero, e.g. reference levels for contaminants sedimentation in soil, water or air; or the distance between the centres (axes) of elements that should be concentric (coaxial). Kotulski and Szczepiński [25] give examples showing a good description of vanadium concentration in sediments and cadmium concentration in soil using the log-normal distribution. The study [28] shows that the log-normal distribution can be appropriate for describing the fragmentation test conducted for fibre-reinforced composites. An example of applying this distribution to describe a parameter in a model of fatigue crack growth can be found in Doliński [12]. Log-normal distribution applications associated with the problems of the reliability theory are offered in the works [23, 10, 28, 33], and for the extreme value theory applied to warning forecasts in hydrology they are found in the work [26]. Basic characteristics of the LN distribution are provided in Table 1 [29].

Table 1. Basic characteristics of the log-normal distribution

Parameters	$0 \leq \mu < \infty, \quad \sigma > 0$
Support	$x \in (0, \infty)$
Expected value	$e^{\mu + \frac{\sigma^2}{2}}$
Variance	$(e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$

It is known that the maximum likelihood estimators of the log-normal distribution parameters are equal to [24]:

$$\hat{\mu} = \frac{\sum_{i=1}^n \ln x_i}{n}, \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (\ln x_i - \hat{\mu})^2}{n}. \quad (3)$$

Let us consider a sample of n observations, which, after arranging in an ascending order, form a sequence: $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$. Therefore, $x_{(i)}$ denotes the value of the i -th positional statistics $X_{i:n}$ from an n -element simple sample x . The following notation have been adopted in the further part of this article:

$\Psi(x, \mu, \sigma^2)$ denotes the probability density function of distribution $LN(\mu, \sigma^2)$,

$\Phi(x, \mu, \sigma^2)$ denotes the cumulative distribution function of $LN(\mu, \sigma^2)$,

$f_{i:n}(x, \mu, \sigma^2)$ denotes the probability density function of the i -th positional statistics $X_{i:n}$.

Therefore, we have [11]:

$$\Psi(x, \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \quad x > 0, \quad (4)$$

$$\Phi(x, \mu, \sigma^2) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left[\frac{\ln x - \mu}{\sigma\sqrt{2}}\right], \quad x > 0, \quad (5)$$

$$f_{i:n}(x, \mu, \sigma^2) = [B(i, n-i+1)]^{-1} [\Phi(x, \mu, \sigma^2)]^{i-1} [1 - \Phi(x, \mu, \sigma^2)]^{n-i} \Psi(x, \mu, \sigma^2), \quad (6)$$

where erf is a special function (non-elementary) called the Gauss error function, $B(p, q)$ denotes the beta special function:

$$B(p, q) = \int_0^1 t^{p-1} (1-t)^{q-1} dt, \quad p > 0, \quad q > 0. \quad (7)$$

It is known that:

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \frac{(p-1)!(q-1)!}{(p+q-1)!} \quad (8)$$

for natural p and q , while $\Gamma(x)$ denotes the Euler's gamma special function.

Denoting $EX = \mu'$, $VAR(X) = (\sigma')^2$ one can observe in the Table 1 that both μ' and $(\sigma')^2$ depend on μ , σ^2 parameters and vice versa. Therefore, all the functions Φ , Ψ and $f_{i:n}$ are dependent on μ' and $(\sigma')^2$.

5. Model of outliers

Consider the following situation: an ordered n -element sample is given:

$$\underbrace{x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n_1)}}_{n_1 \text{ of the lowest observations}} \leq \overbrace{x_{(n_1+1)} \leq \dots \leq x_{(n-n_2)}}^{\text{Main part of the sample}} \leq \underbrace{x_{(n-n_2+1)} \leq \dots \leq x_{(n)}}_{n_2 \text{ of the highest ones}} \quad (9)$$

from the log-normal distribution. The 'main part' of the sample group is derived from a population with an average of μ' , and the outliers may constitute the lowest observations group or the highest observations group. They originate from populations of different μ'_1 and μ'_2 average values. Detection of outliers is generally performed using the procedure of hypothesis testing at a certain level of significance. Then, relevant hypotheses have the following form:

- H_0 – lack of outliers, that is $\mu'_1 = \mu' = \mu'_2$,
- H_{1a} – there exist lowest outliers, that is $\mu'_1 < \mu'$,
- H_{1b} – highest outliers, that is $\mu' < \mu'_2$,
- H_{1c} – there exist lowest and highest outliers, that is $\mu'_1 < \mu' < \mu'_2$,

Here outliers are determined using the Akaike information criterion. Thus, by setting the parameters accordingly: $\mu_1 < \mu < \mu_2$ and the same σ for all groups of observation, we obtain:

$$\mu'_1 < \mu' < \mu'_2 \quad \text{and} \quad \sigma'_1 < \sigma' < \sigma'_2. \quad (10)$$

Therefore, the model with outliers can be described in the following way:

$$h(x_{(i)}) = \begin{cases} \Psi(x_{(i)}, \mu_1, \sigma^2) & \text{for } i = 1, \dots, n_1 \\ f_{i-n_1:n-n_1-n_2}(x_{(i)}, \mu, \sigma^2) & \text{for } i = n_1 + 1, \dots, n - n_2 \\ \Psi(x_{(i)}, \mu_2, \sigma^2) & \text{for } i = n - n_2 + 1, \dots, n \end{cases} \quad (11)$$

6. Akaike information criterion

In order to determine the value of the Akaike information criterion (1) for the model (11) we find its likelihood function:

$$L(n_1, n_2, \mu, \mu_1, \mu_2, \sigma^2 | x) = \prod_i^{n_1} \Psi(x_{(i)}, \mu_1, \sigma^2) \prod_{i=n_1+1}^{n-n_2} f_{i-n_1:n-n_1-n_2}(x_{(i)}, \mu, \sigma^2) \prod_{i=n-n_2+1}^n \Psi(x_{(i)}, \mu_2, \sigma^2)$$

Thus, in view of (4), (5) and (6) the logarithm of the likelihood function equals to:

$$\begin{aligned} \ln L(n_1, n_2, \mu, \mu_1, \mu_2, \sigma^2 | x) &= \ln \prod_{i=1}^{n_1} \frac{1}{x_{(i)}\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x_{(i)} - \mu_1)^2}{2\sigma^2}\right] + \\ &+ \ln \left[\prod_{i=n_1+1}^{n-n_2} [B(i-n_1, n-n_1-n_2-i+n_1+1)]^{-1} [\Phi(x_{(i)}, \mu, \sigma^2)]^{i-n_1-1} [1 - \Phi(x_{(i)}, \mu, \sigma^2)]^{n-n_1-n_2-i+n_1} \times \right. \\ &\left. \times \frac{1}{x_{(i)}\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x_{(i)} - \mu)^2}{2\sigma^2}\right] \right] + \ln \prod_{i=n-n_2+1}^n \frac{1}{x_{(i)}\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x_{(i)} - \mu_2)^2}{2\sigma^2}\right], \end{aligned} \quad (12)$$

that is:

$$\left\{ \begin{aligned} & l(n_1, n_2, \mu, \mu_1, \mu_2, \sigma^2 | x) = -\frac{1}{2} \left\{ n \ln(2\pi) + n \ln(\sigma^2) + \frac{1}{\sigma^2} \sum_{i=1}^n (\ln x_{(i)} - \mu^i)^2 \right\} \\ & - \sum_{i=1}^n \ln x_{(i)} - \sum_{i=n_1+1}^{n-n_2} \left\{ \ln B(j, k-j+1) - (j-1) \ln \Phi(x_{(i)}) - (k-j) \ln [1 - \Phi(x_{(i)})] \right\} \end{aligned} \right\} \quad (13)$$

where $j = i - n_1$, $k = n - n_1 - n_2$ and:

$$\mu^i = \begin{cases} \mu_1 & \text{if } 1 \leq i \leq n_1 \\ \mu & \text{if } n_1 < i \leq n - n_2 \\ \mu_2 & \text{if } n - n_2 < i \leq n \end{cases} \quad (14)$$

Finally, for the observed sample x the value of the Akaike information criterion function $AIC(n_1, n_2)$ dependent on n_1 and n_2 equals:

$$AIC(n_1, n_2) = \begin{cases} -2l(n_1, n_2; \hat{\mu}, \hat{\sigma}^2 | x) + 2 \cdot 2 & \text{if } n_1 = 0, n_2 = 0 \\ -2l(n_1, n_2; \hat{\mu}, \hat{\mu}_1, \hat{\sigma}^2 | x) + 2 \cdot 3 & \text{if } n_1 \neq 0, n_2 = 0 \\ -2l(n_1, n_2; \hat{\mu}, \hat{\mu}_2, \hat{\sigma}^2 | x) + 2 \cdot 3 & \text{if } n_1 = 0, n_2 \neq 0 \\ -2l(n_1, n_2; \hat{\mu}, \hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}^2 | x) + 2 \cdot 4 & \text{if } n_1 \neq 0, n_2 \neq 0 \end{cases} \quad (15)$$

where $\hat{\mu}$, $\hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\sigma}^2$ denote estimators of model parameters obtained with the maximum likelihood method.

7. Experimental research

Experimental research was conducted at the Institute of Transportation Systems and Electrical Engineering of the University of Technology and Humanities in Radom. Driver response times were recorded at a fixed intensity level of real road traffic and fixed biometeorological conditions using a device allowing for measurements in both stationary and dynamic conditions. The device received a signal by radio from the preceding vehicle, in which a transmitter activated by pressing the brake pedal was installed. The voltage powering the brake lights switched on the signal in the measurement impulse transmitter. An impulse receiver and an electronic time measuring device

were installed in the vehicle following with the test driver. This device measured the number of measurement impulses and converted this number into time, with the accuracy of $10^{-3}s$. Pressing the brake pedal by the test driver resulted in stopping the time measuring device and saving the driver response time with the above-mentioned accuracy. At the same time the situation on the road was filmed through the windscreen [23].

8. Results

An example of a representative measurement result stream of the response time obtained from one of the drivers is illustrated in the histogram (Fig. 2). It appears that the driver response time has a log-normal distribution. This is confirmed by the shape of the histogram, and the results of statistical tests (Table 2). The values of all the tests: Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling, presented in the third column of Table 2, clearly indicate that there are no grounds to reject the hypothesis of the log-normal distribution of driver response time. The calculations of tests values and drawings were performed with the SAS version 9.3 program.

Using the theory developed in the previous chapters, the values of the Akaike information criterion AIC for various configurations of outliers (different numbers of the lowest and highest outliers) were found. The results are presented in Table 3.

As presented in Table 3, the lowest value of the Akaike information criterion AIC , marked with an asterisk, indicates that in the studied stream the outlying response times are the ones from the set containing two lowest and six highest values. Highlighted in Figure 3 are the times significantly deviating from a straight line which, using a logarithmic scale, corresponds to the regression line $y = 8.056 \cdot \exp(0.0306 \cdot i)$ for $n_1 < i \leq n - n_2$ which provides characteristics of the psycho-physical condition of the analysed driver within the studied period. After rejecting extreme observations, the values in 2 out of 3 tests checking the fitting of data stream to the log-normal distribution in Table 2,

Table 3. Values of the information criterion AIC for different configurations of outliers

		Number of the highest outlier values										
		0	1	2	3	4	5	6	7	8	9	10
Number of the lowest outlier values	0	-30.10	-41.77	-42.85	-44.95	-46.84	-47.75	-45.28	-40.82	-36.02	-30.40	-24.93
	1	-23.35	-37.29	-39.57	-43.38	-47.37	-50.75	-50.22	-46.33	-42.26	-37.43	-31.70
	2	-16.25	-31.55	-35.10	-40.71	-46.88	-52.49	-53.43*	-50.10	-46.58	-41.80	-35.23
	3	-12.11	-27.34	-30.78	-36.39	-42.69	-48.51	-49.48	-45.86	-42.01	-36.69	-29.35
	4	-7.68	-22.82	-26.19	-31.86	-38.34	-44.44	-45.45	-41.52	-37.29	-31.24	-22.91
	5	-3.64	-18.57	-21.80	-27.43	-33.96	-40.17	-41.02	-36.66	-31.86	-24.73	-15.16
	6	-0.36	-14.98	-17.97	-23.44	-29.87	-35.98	-36.46	-31.55	-26.01	-17.49	-6.53
	7	3.38	-10.84	-13.59	-18.87	-25.17	-31.04	-30.87	-25.19	-18.53	-7.72	5.40
	8	4.07	-9.83	-12.22	-17.11	-23.00	-28.42	-27.62	-21.31	-13.80	-1.53	12.72
	9	4.65	-9.00	-11.06	-15.59	-21.11	-26.13	-24.74	-17.82	-9.49	4.24	19.57
	10	5.48	-7.98	-9.71	-13.92	-19.09	-23.74	-21.74	-14.17	-4.95	10.45	27.00

* optimal configuration of the sample

Table 2. Tests of observation conformance to the log-normal distribution

Test	Data with outliers		Data after rejection of outliers	
	Statistics	p-value	Statistics	p-value
Kolmogorov-Smirnov	$D = 0.1244$	> 0.150	$D = 0.1003$	> 0.150
Cramer-von Mises	$W^2 = 0.1169$	0.067	$W^2 = 0.0412$	0.5
Anderson-Darling	$A^2 = 0.6769$	0.075	$A^2 = 0.3109$	0.5

reach the median value. Therefore, they unequivocally indicate that there are no grounds for rejecting the hypothesis of typical driver response time having a log-normal distribution. In addition, they confirm the validity of identifying those outliers as untypical of this driver. The validity of the chosen methodology is confirmed by the full compatibility of all applied statistical procedures.

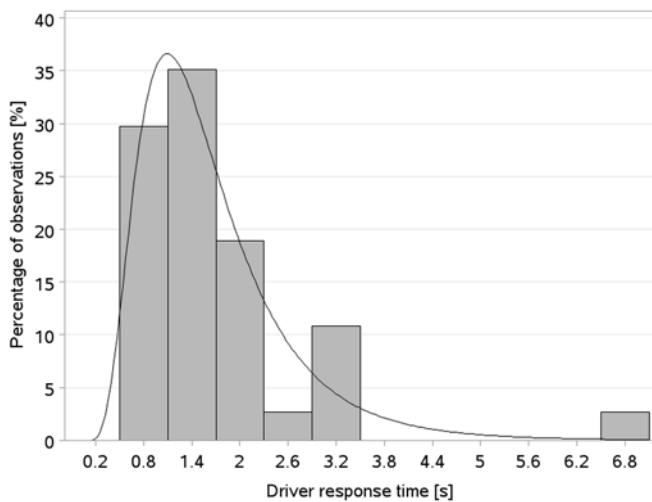


Fig. 2. The distribution histogram of the observed response time stream of one of the drivers, and conforming log-normal distribution.

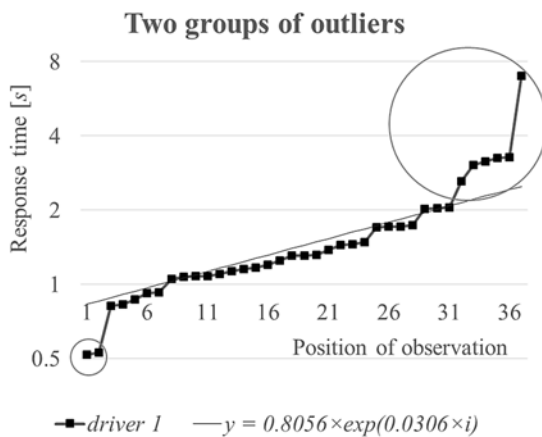


Fig. 3. The two observed groups of driver response time are critical to the reliability level of the driver-vehicle system in the process of driving

9. Conclusions

The probability function $R(t) = 1 - F(t) = P(T \geq t)$ conventionally known as the reliability function (*reliability function*) or survival function, on account of the pejorative nature of the long T time of the driver's response, should rather be termed as the "*risk function*" here. Due to the fact that for $t = 0$ value $R(0) = 1$ and $\lim_{t \rightarrow \infty} R(t) = 0$, the "risk function" is a non-growing function of t time and it expresses the probability of the right-hand "tail" distribution. That is to say, the "risk function" refers to the description of probability for the extreme values being the most adverse to the safety of road traffic. Therefore, attention should be paid to the rate of convergence of the right tail to 0. In systems described by normal distributions, the occurrence of extreme events (from tails) is so low that, in practice, such events are not observed. It is expressed by the three-sigma rule. While a single long driver response time constitutes a major threat to road safety. Hence, the statement that the driver response time in real traffic has a log-normal distribution means that the distribution of the driver response time has a heavy right-hand tail, i.e. the response times can reach relatively high values with significant probabilities. Therefore, the identification of driver response time in real traffic as a variable having a large right-hand tail of a log-normal distribution shows it to be a major risk source in road traffic. And then the observation made in the representative Fig. 3, whereby the six highest outliers from the right-hand tail significantly exceed the values of the already heavy right-hand tail, additionally compounds the horror of dangerous situations. Moreover, the downward deviation of the two shortest response times from the typical time distribution presented in Fig. 3, points to the driver's nervous reactions. These results may also indicate loss of (deterioration of) driving smoothness, and this is correlated with a risk increase, and consequently with the risk of traffic accident occurrence. Therefore, the additional identification of both groups of critical driver response times with a heavy right-hand tail of their distribution combines to indicate the existence of three potential threats to the reliability of the driver-vehicle system on the part of the driver, resulting in the deterioration of road safety. Thus, it is necessary to seek systems of driver support which are aimed at eliminating these threats.

References

1. Akaike H. Information theory and an extension of the maximum likelihood principle. 2nd International Symposium on Information Theory. Eds. B.N. Petru and F. Csaki. Budapest: Akademia Kiado, 1973; 267-281.
2. Akaike H. On entropy maximization principle. Proc Symposium on Applications of Statistics. Ed. P.R. Krishnaiah. Amsterdam: North Holland, 1977; 27-47.
3. Archer J. Methods for the Assessment and Prediction of Traffic Safety at Urban Intersections and their Application in Micro-simulation Modelling Academic Thesis. Division of Transport and Logistics, Stockholm: Royal Institute of Technology, Kungliga Tekniska Högskolan (KTH) SE-100 44, 2004.
4. Ashton W D. The theory of road traffic flow. London: Methuen & Co. Ltd. 1966.
5. Badger J E. Human factors affecting perception. Law and Order Magazine, 1996.
6. Barnett V, Lewis T. Outliers in Statistical Data. John Wiley & Sons, 1994.
7. Benjamin J R, Cornell C A. Probability, Statistics, and Decision for Civil Engineers. New York: McGraw-Hill Book Company, 2014.
8. Breuning M, Kriegel H P, Sander J. LOF: Identifying Density-Based Local Outliers. In Proceedings of the ACM SIGMOD Conference, 2000; 93-104.
9. Cieślak K, Karpińska O. Przyszłość za kierownicą – czy samochód będzie myśleć za nas? Translogistics 2014; 95-108.
10. Cooke R M, Nieboer D. Heavy-tailed distributions: Data, diagnostics, and new developments. Resources for the Future Discussion Paper, Washington, 2011; (11-19), [online: <http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-DP-11-19.pdf>, access: 2016-04-03].
11. David H A, Nagaraja H N. Orders Statistics. Wiley Series in Probability and Statistics, 2003, <https://doi.org/10.1002/0471722162>.
12. Doliński K. Comparison of stochastic model of fatigue crack growth with experiments. Fatigue Fract. Engng. Mater. Struct. 1993; 16 (10): 1021-1034, <https://doi.org/10.1111/j.1460-2695.1993.tb00075.x>.
13. Drew D R. Traffic flow theory and control. New York: McGraw-Hill Book Company, 1968; Chapter 12.
14. Euro RAP [online: <http://www.eurorap.org/>, access: 2016-04-03].
15. Ferguson T S. On the rejection of outliers. In Proc. Fourth Berkeley Symposium Math. Statist. Prob. 1961; 1: 253-287.

16. Foss S, Korshunov D, Zachary S. An Introduction to Heavy-Tailed and Sub exponential. Series in Operations Research and Financial Engineering Distributions. Springer, 2013, <https://doi.org/10.1007/978-1-4614-7101-1>.
17. Fronczak A, Fronczak P. Świat sieci złożonych. Od fizyki do Internetu. Warszawa: Wydawnictwo Naukowe PWN, 2009.
18. Golab L, Ozsu M T. Issues in data stream management. ACM Sigmod Record, 2003; 32 (2): 5-14, <https://doi.org/10.1145/776985.776986>.
19. Grubbs F E. Sample criteria for testing outlying observations. Ann. Math. Statist. 1960; 1: 27-58.
20. Guzowski P. Wypadki drogowe – statystyka i przyczyny. Szkolenie z zakresu ratownictwa technicznego dla Strażaków Ratownictwa OSP. Red. Surala Zb. CNBOP. 2008; 4-23.
21. Jones T R. and Potts R B. The measurement of acceleration noise – a traffic parameter. Operations Research, 1962; 10: 745-763, <https://doi.org/10.1287/opre.10.6.745>.
22. Jurecki R, Jaśkiewicz M, Guzek M, Lozia Z, Zdanowicz P. Driver's reaction time under emergency braking a car – Research in a driving simulator. Eksploatacja i Niezawodność – Maintenance and Reliability. 2012; 14 (4): 295–301.
23. Kaczor M, Marczak M. (red.) Badanie wpływu oznakowania tyłu pojazdu na jakość ruchu. – Praca naukowo badawcza. - Etap II. Zakład Modelowania Procesów Transportowych w Instytucie Systemów Transportowych i Eksploatacji Taboru Kolejowego, WSI Radom, 1983.
24. Krzyśko M. Mathematical Statistics. Poznań: Wydawnictwo Naukowe UAM, 2004 (in Polish).
25. Kotulski Z, Szczepiński W. Rachunek błędów dla inżynierów. Warszawa: WNT, 2004.
26. Kuźmiński Ł. The application of value extreme theory in warning forecast in hydrology for sequence of independent lognormal random variables. Studia Ekonomiczne, 2014; 207: 148-158 (in Polish).
27. Mercedes-Benz "Intelligent Drive". TecDay. Press Information. November,
28. Murzewski J. Niezawodność konstrukcji inżynierskich. Warszawa: Arkady, 1989.
29. Patel J, Kapadia C H, Owen D B. Handbook of statistical distribution. Marcel Dekker, 1976.
30. Rousseeuw P, Leroy A. Robust Regression and Outlier Detection. John Wiley & Sons, 2003.
31. Sakamoto Y, Ishiguro M, Kitagawa G. Akaike Information Criterion Statistics. Tokyo: Reidel Publishing Company, 1986.
32. Srivastava M S, Von Rosen D. Outliers in Multivariate Regression Models. J. Mult. Anal. 1998; 65: 195-208, <https://doi.org/10.1006/jmva.1997.1729>.
33. Suchodolski S. Probabilistyczna ocena wytrzymałości włókien szklanych oraz połączenia włókno-spoivo epoksydowe na podstawie testu fragmentacji. Prace XVII Sympozjum Mechaniki Eksperymentalnej Ciała Stałego, Jachranka 1996, 550-557.
34. Wicher J. Bezpieczeństwo samochodów i ruchu drogowego. Warszawa: Wydawnictwa Komunikacji i Łączności, 2012.
35. Woch J. Teoria potoków ruchu. Katowice: Zakład Inżynierii Ruchu Instytutu Transportu, 2001 [online: http://dydaktyka.polsl.pl/kir/TPR_WWW/TPR.htm, access: 2016-04-03].

Andrzej KORNACKI**Jacek WAWRZOSEK****Andrzej BOCHNIAK**

Department of Applied Mathematics and Computer Science

University of Life Sciences in Lublin

ul. Głęboka 28, 20-612 Lublin, Poland

E-mails: andrzej.kornacki@up.lublin.pl, jacek.wawrzosek@up.lublin.pl, andrzej.bochniak@up.lublin.pl**Andrzej SZYMANEK**

Institute of Transportation Systems and Electrical Engineering

University of Technology and Humanities in Radom

ul. Malczewskiego 29, 26-600 Radom, Poland

E-mail: a.szzymanek@uthrad.pl**Halina PAWLAK**

Department of Technology, Ergonomics Unit

University of Life Sciences in Lublin

ul. Głęboka 28, 20-612 Lublin, Poland

E-mail: halina.pawlak@up.lublin.pl