www.czasopisma.pan.pl *Metrol. Meas. Syst.*, Vol. 25 (2018) No. 4, pp. 743–754. www.journals.pan.pl

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

ACCURACY CRITERIA FOR EVALUATION OF WEIGH-IN-MOTION SYSTEMS

Piotr Burnos, Janusz Gajda, Ryszard Sroka

AGH University of Science and Technology, Department of Measurement and Electronics, A. Mickiewicza 30, Cracow 30-059, Poland (*burnos@agh.edu.pl*, +48 12 633 8565, *jgajda@agh.edu.pl*, *rysieks@agh.edu.pl*)

Abstract

Measurement data obtained from Weigh-in-Motion systems support protection of road pavements from the adverse phenomenon of vehicle overloading. For this protection to be effective, WIM systems must be accurate and obtain a certificate of metrological legalization. Unfortunately there is no legal standard for accuracy assessment of Weigh-in-Motion (WIM) systems. Due to the international range of road transport, it is necessary to standardize methods and criteria applied for assessing such systems' accuracy. In our paper we present two methods of determining accuracy of WIM systems. Both are based on the population of weighing errors determined experimentally during system testing. The first method is called a reliability characteristic and was developed by the authors. The second method is based on determining boundaries of the tolerance interval for weighing errors. Properties of both methods were assessed on the basis of simulation studies as well as experimental results obtained from a 16-sensor WIM system.

Keywords: Weigh-in-Motion systems, enforcement of vehicles, accuracy assessment, tolerance intervals, reliability characteristic.

1. Introduction

It is proved that the damaging effect of one heavy vehicle, *i.e.* a vehicle with GVM greater than 30 t, on a road pavement is equal to the impact caused by tens of thousands of passenger cars [1]. Moreover, an overloaded heavy vehicle causes several times greater fatigue damage to a pavement structure than a properly loaded one. In the last 40 years extensive studies concerning this problem were carried out [2–5]. Some tools used for detecting overloaded vehicles and measuring their weight and axle loads are *Weigh-in-Motion* (WIM) systems. The idea of such systems consists in measuring dynamic loads of a moving vehicle's wheels on a road pavement. On that basis values of static load and *gross vehicle weight* (GVW) are estimated, which are used as the overloading criteria [6–7]. Fig. 1 presents an example of WIM system.

The inaccuracy of the mass measurement, results in a necessity of increasing the permissible mass value specified in the applicable regulations, by the maximum value of error reported by WIM system. Such treatment is necessary due to the caution required. It provides a safety margin enabling to avoid erroneous recognition of a normative vehicle as an overloaded one. As a result, some of the vehicles that are actually overloaded will not be eliminated from traffic. Thus, the

Article history: received on Feb. 28, 2018; accepted on Aug. 14, 2018; available online on Dec. 31, 2018, DOI: 10.24425/mms.2018.124881.

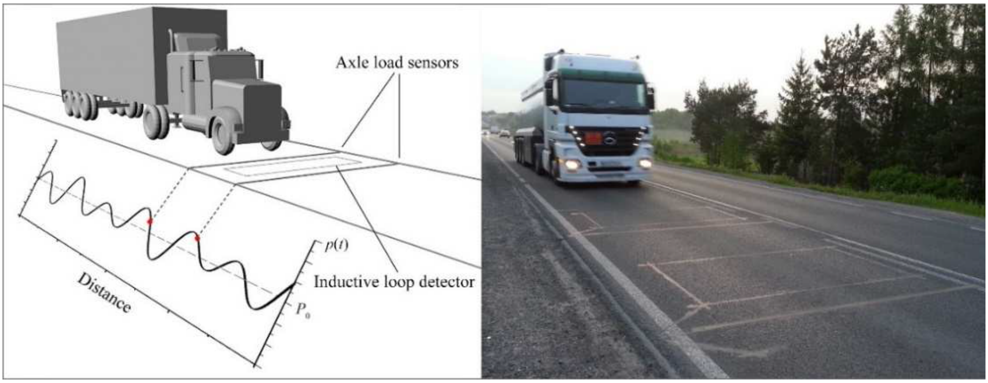

Fig. 1. An example of Weigh-in-Motion system.

effectiveness of WIM system is limited. Therefore, a proper estimation of weighing error is crucial for WIM systems to be efficient in direct enforcement of overloading.

The weighing accuracy is influenced by different factors. A dominant role is played by the fact that due to the vehicle motion, the sensors respond to the dynamic component $p(t)$ of a vehicle axle load. Hence, the weighing accuracy is affected by the vehicle speed, the mechanical properties of its suspension and the quality of road pavement where WIM system is installed [8, 9]. Other factors that influence the system accuracy are properties of the applied load sensors, in particular the non-uniform distribution of sensor sensitivity as a function of its length. Also, frequency of the system calibration affects the accuracy of weighing results. A description of factors influencing the accuracy of WIMs can be find in our paper [10]. The accuracy of WIM systems depends also on the algorithms employed for the static load and GVW estimation [6], as well as on the climatic conditions in which the WIM system is operated. The influence of the pavement temperature is of particular significance for embedded sensors because it is a source of systematic error (bias) in weighing results. A detailed analysis of this phenomenon is presented in our works [11, 12].

The fact that measurements are made in dynamic conditions leads to many additional problems associated with determining a reference value for the calibration procedure and developing methods for WIM system accuracy assessment.

In this paper we present our approach to determining the accuracy criteria of WIM systems based on the population of weighing errors determined during system testing. The first method is related to the determination of WIM system reliability characteristic while the second one – to the determination of tolerance intervals for weighing errors. Both methods:

– have theoretical foundations;
– take into account both components of errors: random and systematic;
– are simple – no difficult numerical calculations are needed;

and in the case of reliability characteristic no assumptions concerning the distribution class of WIM system error population is needed.

Properties of both methods were assessed on the basis of the simulation study results and the experimental results obtained from a 16-load sensor WIM system, built by the authors of this work.

2. Review of existing accuracy assessment methods for WIM systems

The method of pre-weighed vehicles is often used to assess the accuracy of WIM systems. This method uses several vehicles pre-weighed on a static scales, which repeatedly pass through the tested WIM system with different speeds. The subject of measurement on a static scales is GVM and/or static load of individual axles, which are used as the reference quantity.

From the measurement results obtained by this method, a set of relative errors is calculated according to the relation (1):

$$\delta_i = \frac{w_i - w_i^{ref}}{w_i^{ref}}, \tag{1}$$

where: w_i – a weighing result of the i-th pre-weighted vehicle obtained from the tested WIM system; w_i^{ref} – a reference result of weighing the i-th vehicle pre-weighted on a static scales.

Several methods for WIM system accuracy assessment have been proposed in the literature and put into practice. All methods focus on the analysis of standard deviation of the set (1) and determining the random component of error. The systematic component, a so called bias error, is not taken into account. In this section we shortly review all methods.

The method described in [13] by Slavik is based on computer simulations. The statistical parameters of error are determined from a limited number of measurements made at a WIM site for consecutive passes of pre-weighed vehicles. This population is enlarged by means of computer simulations. Pseudo-random samples with the same statistical parameters as those of the experimentally determined system errors are generated in the computer. The basic drawbacks of this method are:

– the use of synthetic instead of experimental data;
– the assumption of a normal distribution of errors (1) which is difficult to verifying when the population of these errors is too scarce;
– the substitution of the probability of occurring an error with an a priori determined value for the measured error.

The method proposed by the *American Society for Testing and Materials Standard* (ASTM) [14] is based on measurement data acquired during WIM system calibration. The accuracy of WIM system is assessed by the estimation of the probability of error exceeding an arbitrarily assumed allowable value. If this probability exceeds 0.05 the system is regarded as insufficiently accurate. In this method the validation of WIM system is based on errors observed during the system calibration, not on those which occur during its operation. However, the probability of exceeding the allowable error value by different systems can be comparable.

The method described in [6] has become an unofficial standard in Europe. However, in relation to it, various objections have been formulated. First of all, they concern a necessity of arbitrary assessment of so-called "repeatability of test conditions", which characterise the environmental conditions the test of WIM system is performed in. Depending on the reproducibility degree and the size of the set of vehicles used in the test, the minimum value of the confidence level is arbitrarily assumed. Exceeding this value means malfunctioning of the tested WIM system. Thus, the selected "repeatability of test conditions" and confidence level correspond directly to the accuracy assessment result.

The method described in [15] is commonly accepted in Australia, where generally no obligatory standards exist. The described method is simple and consists in determining a limit value of the weighing error (of a single axle load, axle group load, or GVW) that will not be exceeded by at least 95% of weighed vehicles. Assuming a normal distribution of errors, this is equivalent to taking the confidence interval width equal to twice the standard deviation. In practice, the standard deviation is computed, often without any verification of normality of the error distribution, and the system error is assumed to be twice the standard deviation value. Similar limitations as in the case of the method recommended by ASTM occur in this method.

In all described methods it is assumed that a systematic error (bias) is eliminated in WIM system and only a random component of error is evaluated. In general it is not true. As we showed in our paper [12], due to disturbances, *e.g.* temperature changes in the system, a systematic error may occur in weighing results. In such a case the accuracy assessment of WIM system requires taking into account a bias error. Both proposed by us methods fulfil this requirement.

3. New approach – reliability characteristic

The reliability characteristic for WIM system accuracy assessment was developed by the authors. The method uses several pre-weighed vehicles (with GVW selected in such a way as to cover uniformly the tested system measurement range) repeatedly passing through the tested WIM system with different speeds. From the obtained measurement results a set of relative errors, computed according to the relation (2), is determined. No assumption regarding error distribution is needed:

$$\delta_i^{abs} = |\delta_i| = \left| \frac{w_i - w_i^{ref}}{w_i^{ref}} \right| . \tag{2}$$

Our approach is based on the statistical analysis of the experimentally determined set of errors (2) using the characteristic (3) for this purpose:

$$\Phi\left(\delta^{abs}\right) = 1 - P\left(\delta^{abs}\right), \tag{3}$$

where: $P(\delta^{abs})$ – a cumulative distribution function; δ^{abs} – a random variable; which values constitute modules of the relative error values (2).

The characteristic (3) is called a system reliability characteristic and determines the probability of the occurrence of a weighing error with a value greater than δ^{abs}. Thus, on this characteristic we can also distinguish an error $\delta_{0.95}$ with a value corresponding to the probability of its occurrence $P = 0.05$. This means that the error $\delta_{0.95}$ is such a value of the function argument (3), for which the probability takes a value of 0.05, *i.e.* $\Phi(\delta_{0.95}) = 0.05$.

The function (3) provides comprehensive information about the system accuracy. For example, in Fig. 2 two characteristics are given with the same probability value for $\delta_{0.95} = 0.1$.

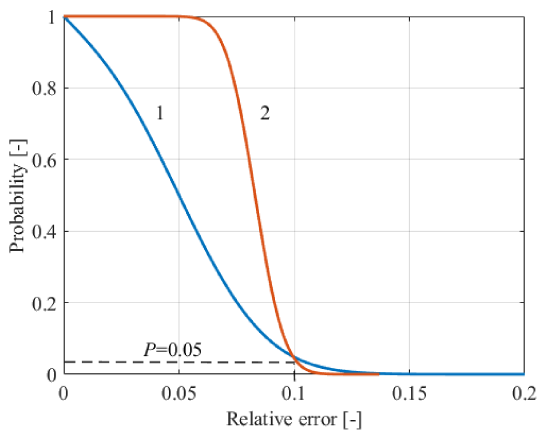

Fig. 2. Examples of reliability characteristics (3).

Distinguishing these systems based only on the standard deviation or error $\delta_{0.95}$ value would be impossible. Taking into account the shape of the characteristics, one can easily distinguish these two systems and notice that system "1" has a better accuracy than system "2". No necessity of assuming the error population distribution is an advantage of this method.

It is obvious that the error value $\delta_{0.95}$ determined on the basis of the reliability characteristic is the same as the error value determined on the basis of the probability density function. The reliability characteristic can be directly estimated on the basis of error values (2) obtained during WIM system testing.

4. Tolerance interval

In all of the previously described methods the accuracy assessment is based on the results obtained during testing of WIM system. However, the accuracy of current measurement results supplied by the system during its normal operation is essential for the system user. Thus, the task should be formulated as follows: on the basis of the results gathered during the WIM system testing, the boundaries of the statistic interval (within which will be placed errors of a given part of the future weighing results, *e.g.* 0.95) should be determined.

The solution for such a problem, for a random variable with a normal distribution, was developed by Wilks in 1941 [16] and Proschan in 1953 [17]. The interval, fulfilling such formulated expectations of the WIM system user, is called a tolerance interval and its boundaries are determined with the dependence (4) for the probability $p = (1-\alpha)$:

$$\delta^{\pm}_{(1-\alpha)} = \hat{\mu} \pm t^{(N-1)}_{(1-\alpha/2)} \sqrt{\frac{N+1}{N}}\,\hat{\sigma}, \tag{4}$$

where: $t^{(N-1)}_{(1-\alpha/2)}$ – student's t distribution variable with $(N-1)$ degrees of freedom, determined for the probability $(1-\alpha/2)$; N – population size *i.e.* the number of measurement results (errors) obtained during the WIM system test; $\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}\delta_i$ – estimate of the expected value; $\hat{\sigma} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\delta_i - \hat{\mu})^2}$ – estimate of the standard deviation; δ_i for $i = 1, 2, \ldots, N$ – error values (1) gathered during the system testing.

The boundaries of the tolerance interval (4) are determined on the basis of a finite number of the observed elements of a general population. In consequence, the dependence (4) only enables these boundaries to be estimated (in a statistic sense). This means that, due to multiple repetitions of the whole procedure, a set of values of random variable boundaries (4) will be obtained. This variability causes that the probability p, that an element of the general population will fall into this interval, also varies. Consequently, it can happen that while looking for boundaries of the tolerance interval which corresponds to a probability $p = 0.95$, we will determine the interval for $p = 0.909$ or $p = 0.98$. The size of sample population has a decisive effect on the uncertainty of the tolerance interval boundaries and thus on the variability of probability p. The rule of thumb holds here "the more elements of sample population the lower random variability of p". In relation to WIM systems the sample population size corresponds to the number of pre-weighted vehicle passes during system testing. In consideration of practical reasons, this size is limited and contained within a range from a dozen or so to several dozens. Therefore, an uncertainty in determining the boundaries of the tolerance interval may be observed.

In practice, the application of the tolerance interval for expressing the uncertainty of weighing results obtained from WIM system is complicated, since in the case when the tolerance interval is not symmetric with reference to zero, it is defined by two numbers $\delta^{\pm}_{(1-\alpha)}$ differing in modulus. Thus, the uncertainty of weighing results obtained from WIM system can be assessed by taking into account the maximum values of modules of both boundaries of the tolerance interval, marked as:

$$\delta^{\max}_{(1-\alpha)} = \max\left|\delta^{\pm}_{(1-\alpha)}\right|. \tag{5}$$

However, with a limited number of pre-weighted vehicle passes, this parameter also indicates a random variability.

5. Expanded tolerance interval

Underestimating the error value is unacceptable in WIM systems for direct enforcement. In the case of using these results for assessing administrative WIM systems, an over-optimistic estimation of their accuracy is especially dangerous. It could cause a standard vehicle to be considered as an overloaded one, thus assuming more cautious accuracy estimations is justified. Therefore, the accuracy of WIM system should be assessed based on the estimation of the maximum error value. For this purpose, an extended tolerance interval can be used (6):

$$\delta^{Ext\pm}_{(1-\alpha)} = \hat{\mu} \pm k\Delta\hat{\mu} \pm t^{(N-1)}_{(1-\alpha/2)}\sqrt{\frac{N+1}{N}}\left(\hat{\sigma} + k\Delta\hat{\sigma}\right), \tag{6}$$

where: k – an arbitrarily assumed expansion coefficient; $\Delta\hat{\mu} = \hat{\sigma}/\sqrt{N}$ – an uncertainty of the estimate of the expected value, determined on the basis of the test of the set size N; $\Delta\hat{\sigma} = \hat{\sigma}/\sqrt{2N}$ – an uncertainty of the estimate of the standard deviation, determined on the basis of the test of the sample set size N.

The assumed value of the expansion coefficient k depends on a cautious level of safety margin.

As before, the uncertainty of weighing results can be assessed by taking into account the maximum values of modules of both boundaries of the extended tolerance interval (6), marked as:

$$\delta^{Ext\max}_{(1-\alpha)} = \max\left|\delta^{Ext\pm}_{(1-\alpha)}\right|. \tag{7}$$

6. Methodology and simulation results

All simulation tests were carried out in Matlab environment with the use of Statistics and Machine Learning Toolbox, but all programs and algorithms were written by the authors. To simplify the interpretation of the results, simulation tests were carried out for an error population (1) with a normal distribution. The most general case was considered, with a negative systematic error (bias) and a random variability of measurement results. Therefore, it was assumed that the expected value of error population is $\mu = -0.15$, and the standard deviation is $\sigma = 0.1$. This results in an asymmetric probability distribution with reference to zero. In addition, it was assumed that the size of error population is $N = 30$, which is a typical value in the case of testing WIM system using pre-weighed vehicles. The value of expansion coefficient k was equal to 2. For the assumed error distribution, the true values of $\delta_{0.95}$ and tolerance intervals are:

$$\delta_{0.95} = 0.31,$$

$$\delta_{(1-\alpha)}^{\max} = 0.35,$$
$$\delta_{(1-\alpha)}^{Ext\max} = 0.44.$$

Figure 3 shows:
– a probability density function (pdf) of error population (1);
– a probability density function abs(pdf) of absolute value of error population (2);
– a reliability characteristic.
The true values of errors are marked with vertical bars.

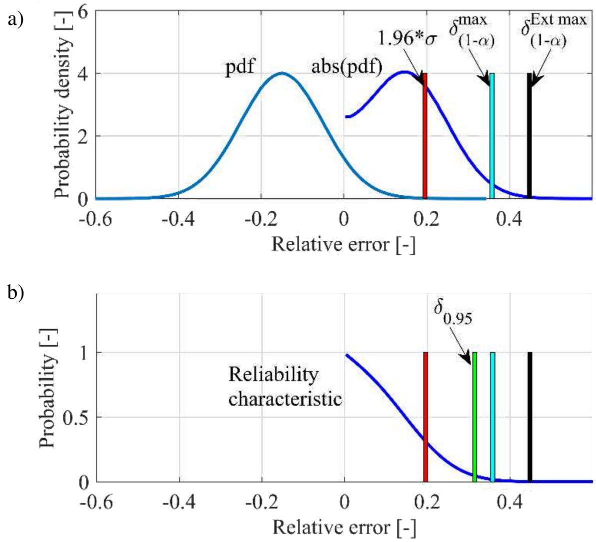

Fig. 3. a) Probability density functions of error populations (1) and (2); b) a reliability characteristic.

For the considered, most general case with a negative systematic error ($\mu = 0.15$) and a small population of weighing errors ($N = 30$) a few conclusions can be formulated based on Fig. 3:
– Reliability characteristics provide comprehensive information on the system accuracy and determine the probability of the occurrence of a weighing error with a value greater than δ^{abs}.
– Double standard deviation 1.96σ has the smallest value in relation to other measures of system error. For this reason, the use of 1.96σ to assess the accuracy of WIM systems for direct enforcement increases the probability of an erroneous interpretation of the weighing results. Underestimating the weighing error causes considering a normative vehicle as an overloaded one.
– For a common size of error population $N = 30$, the error $\delta_{0.95}$ has a higher value than 1.96σ. Thus $\delta_{0.95}$ is a more cautious estimate of the system accuracy.
– The tolerance interval (5) gives an even more cautious estimation of the system accuracy. Its value is greater than the error $\delta_{0.95}$.
– The expanded tolerance interval (7) is a measure of error with the highest value. This is due to the applied expansion coefficient k. Thus, this measure reduces the probability of an erroneous interpretation of weighing results of a vehicle.

– The reliability characteristic and error $\delta_{0.95}$ are determined on the basis of module of error population (2). These estimators take into account both components of error: systematic (bias) and random. In contrast to the measures based solely on the analysis of the standard deviation this is an advantage of the presented methods. What is important, this will not cause underestimating of the WIM system uncertainty. Such an approach seems to be justified in WIM systems for direct enforcement of overloading.

The errors 1.96σ and $\delta_{0.95}$ do not depend directly on the population size N of WIM system errors but they depend on whether the error distribution is symmetrical with reference to zero. In turn, the sample size N affects the tolerance interval and the extended tolerance interval estimators. To illustrate this relationship, a simulation was carried out for various sizes N of population of weighing errors (1) and for two cases of error distribution: a) not symmetrical ($\mu = -0.15$) and b) symmetrical ($\mu = 0$) with reference to zero. In order to reduce the random variability of results, for each N value the calculations were repeated 1000 times, and the results were averaged. Fig. 4 shows values of estimator $\delta_{0.95}$ and estimators (5) and (7) as functions of a population size N of the measurement errors. The true values of 1.96σ and $\delta_{0.95}$ are marked with horizontal lines.

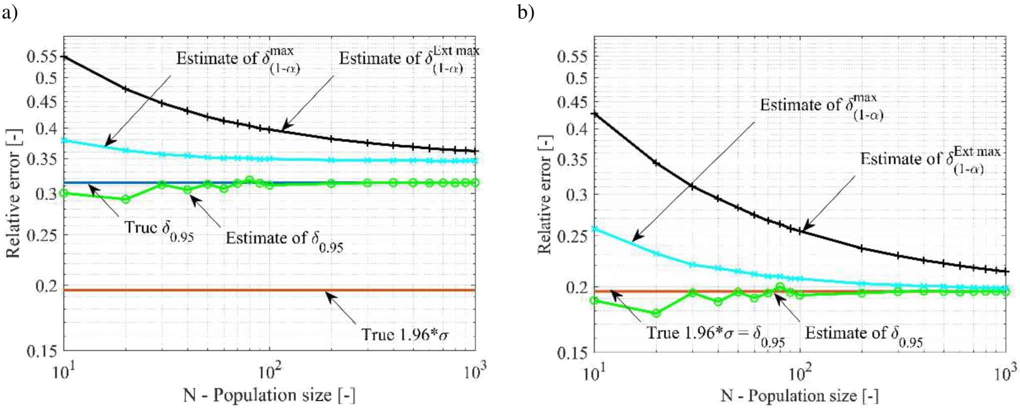

Fig. 4. Values of estimator $\delta_{0.95}$ and estimators (5) and (7) as functions of a population size N of the measurement errors for: a) not symmetrical error distribution ($\mu = -0.15$); b) symmetrical error distribution ($\mu = 0$).

For the most general case with a negative systematic error $\mu = -0.15$ (asymmetric distribution of errors with reference to zero) the following conclusions can be drawn from Fig. 4a:

– $\delta_{0.95}$ error has always a higher value than 1.96σ and gives a more cautious estimate of the system accuracy than double standard deviation no matter the value of the population size N.
– Estimates of tolerance and expanded tolerance intervals give even more cautious estimate of the system accuracy than $\delta_{0.95}$, especially for small N. This is a desired feature in WIM systems for direct enforcement.

For the case without systematic error $\mu = 0$ (a symmetric distribution of errors with reference to zero) the following conclusions can be drawn from Fig. 4b:

– This is a specific case where there is no bias error in WIM system and $\delta_{0.95} = 1.96\sigma$.
– Estimates of tolerance and expanded tolerance intervals asymptotically converge to the value of $\delta_{0.95}$.
– For $N < 100$, which is a common situation, estimation of tolerance intervals gives more cautious estimate of the system accuracy than $\delta_{0.95}$ or 1.96σ. This is a desired feature in

WIM systems for direct enforcement and prevents a normative vehicle from being considered as an overloaded one.

In both cases estimators $\delta^{\max}_{(1-\alpha)}$ and $\delta^{Ext\max}_{(1-\alpha)}$ give more cautious estimate of the system accuracy than $\delta_{0.95}$ or 1.96σ, no matter the value of the population size N. The relationship between measures of errors can be formulated as: $1.96\sigma \le \delta_{0.95} < \delta^{\max}_{(1-\alpha)} < \delta^{Ext\max}_{(1-\alpha)}$. In a general case, when bias error exists in WIM system, double standard deviation should not be used as a measure of WIM system accuracy.

7. Case study

The presented methods for assessment of WIM system accuracy have been compared using data from a real Multi-Sensor Weigh-in-Motion system. The system has been developed by the authors and installed in a site on a national road DK 81 in south of Poland. The site was equipped with 16 piezo-polymer load sensors, eight inductive loop detectors and eight temperature sensors. The applied load sensors use the piezoelectric effect in a polymer called *polyvinylidene fluoride* (PVDF). There were used sensors of Measurement Specialties built in the form of flat belts mounted under the road surface. Therefore, vehicle wheels do not have direct contact with a sensor, and the force signal is transmitted through the road surface. This results in a great sensitivity of measurement results to changes in the surface properties under influence of temperature fluctuations. As a consequence, it results in a lower accuracy of results obtained in WIM systems with sensors of this type. Despite their poor metrological properties, thanks to their moderate price, they are a reasonable alternative to expensive quartz or load cell sensors, especially in multi-sensor systems.

In our system the load sensors are evenly distributed along the WIM site with 1m spacing. Each pair of piezoelectric sensors is associated with one inductive loop detector. This way there are created eight two-sensor configurations, each one as in a classic two-sensor WIM system. This approach enables to achieve a modular structure of the signal conditioning system comprising eight subsystems and eight independent signal processing paths.

The advantages of Multi-Sensor WIM systems result first of all from a higher number of weighing results for each vehicle axle. As a result, the disturbing phenomenon caused by vertical balancing of a weighed vehicle is averaged more effectively. Secondly, a weighted vehicle passing over successive load sensors leads to averaging the effect of non-uniformity of their sensitivity. A drawback of MS-WIM solution (in comparison with a classical WIM system) is its considerable increased construction cost, proportional to the number of installed load sensors, as well as an increased length of WIM site.

The case study was based on weighing results of a) 30 pre-weighed vehicles (a typical number); b) 1029 pre-weighed vehicles (the research experiment). For each set of measurement results the relative errors (1) and their modules (2) were determined. The value of variable χ^2, constructed for the verification of the hypothesis on the distribution normality of the tested WIM system errors, was assumed to be 48.63. For 40 degrees of freedom of this variable and a significance level $\alpha = 0.01$ the critical value was: $\chi^2_{\text{crit}} = 63.691$. Thus, there are no reasons to reject such a hypothesis. Thereby, applying the dependences (4) to (7) to determining the tolerance intervals is justified.

On this basis, the probability density function, reliability characteristic, $\delta_{0.95}$ and tolerance intervals were estimated. The results for a) $N = 30$ and b) $N = 1029$ are presented in Fig. 5.

The estimated values of errors are presented in Table 1.

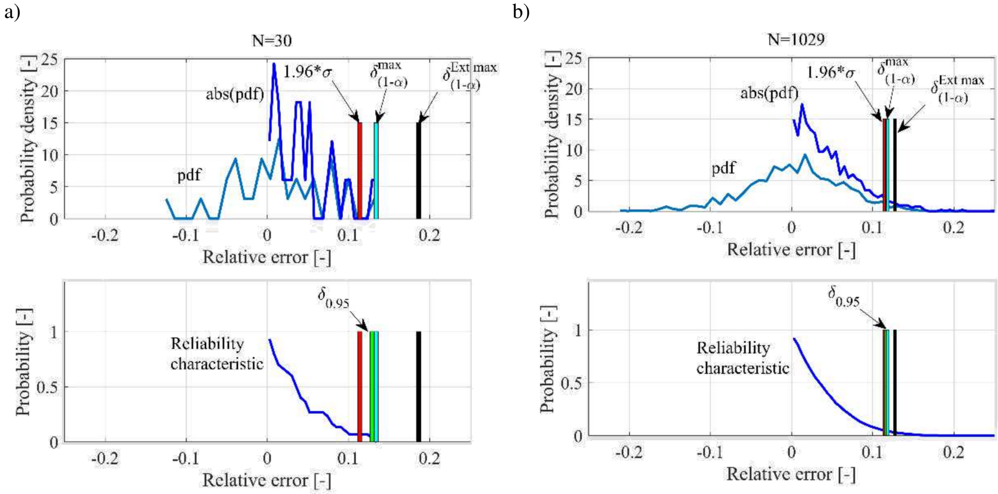

Fig. 5. Probability density functions and reliability characteristics of a real WIM system for: a) $N = 30$; b) $N = 1029$.

Table 1. Estimated values of errors.

	N [–]	
	30	**1029**
1.96σ [–]	0.114	0.113
$\delta_{0.95}$ [–]	0.129	0.116
$\delta_{(1-\alpha)}^{\max}$ [–]	0.134	0.118
$\delta_{(1-\alpha)}^{Ext\max}$ [–]	0.186	0.127

On the basis of the obtained results, the following conclusions can be formulated:

– A low accuracy results to a large extent from the quality of sensors used in our system.
– The simulation results are confirmed by the results obtained for a real WIM system.
– The error population size N influences the estimated values of tolerance interval and expanded tolerance interval. The smaller the population N the more cautious estimate of the system accuracy.
– The relationship observed in the simulations is maintained: $1.96\sigma \leq \delta_{0.95} \leq \delta_{(1-\alpha)}^{\max} \leq \delta_{(1-\alpha)}^{Ext\max}$.
– The error measure $\delta_{(1-\alpha)}^{Ext\max}$ is the most cautious estimation of the WIM system error, close to the maximum error value.

8. Conclusions and future works

The correct estimation of error is very important in WIM systems for direct enforcement because the allowable values of GVW and axle loads must be increased by the amount of this error to achieve a reasonable safety margin. This action results from caution and prevents a normative vehicle from being considered as an overloaded one.

Two ways of assessing accuracy of WIM systems are compared in this paper. The first one is based on the reliability characteristic, whereas the second – on determining the boundaries of the tolerance interval. Conclusions resulting from comparison of the methods are presented below.

- The reliability characteristic does not require any assumptions concerning the distribution class of WIM system error population.
- The error $\delta_{0.95}$ is such a value of the reliability function argument, for which the probability takes a value of 0.05, *i.e.* $\Phi(\delta_{0.95}) = 0.05$.
- The reliability characteristic enables to compare two WIM systems with the same error value $\delta_{0.95}$, and also enables to find the maximum value of this system error.
- Tolerance intervals, single- and double-sided, are based on appropriate theoretical fundamentals and are tools known from many years.
- Using a tolerance interval leads to expressing the WIM system accuracy with two numbers, *i.e.* the lower and upper boundaries of this interval. In consequence, it would be difficult on this basis to compare two different systems.
- Selecting the maximum value out of the tolerance interval boundaries is justified by caution required in the case of administrative WIM systems used for direct enforcement.
- Using tolerance intervals requires knowledge of the distribution of system errors.
- The reliability characteristic, error $\delta_{0.95}$, and estimators of the maximum value of tolerance intervals take into account both systematic (bias) and random error components, in contrast to measures based solely on analysis of the standard deviation of error.
- In general the relationship between measures of errors can be formulated as: $1.96\sigma \leq \delta_{0.95} < \delta^{\max}_{(1-\alpha)} < \delta^{Ext\max}_{(1-\alpha)}$. When bias error exists in WIM system, double standard deviation should not be used as a measure of WIM system accuracy. Tolerance intervals give the most cautious estimate of system accuracy.

The method of classification of WIM systems based on the proposed criteria remains to be developed in the future.

References

[1] Raab, C., Partl, M., Partl, A. (2005). Monitoring traffic loads and pavement deformations on a Swiss motorway. *Post-proceedings of the Fourth International Conference on Weigh-in-Motion*, 301–311.

[2] Rys, D., Judycki, J., Jaskula, P. (2015). Analysis of effect of overloaded vehicles on fatigue life of flexible pavements based on weigh in motion (WIM) data. *Int. J. Pavement Eng.*, 1–11.

[3] Kulakowski, B. (1994). *Vehicle-road interaction*. Philadelphia: ASTM International.

[4] Sadeghi, J.M., Fathali, M. (2007). Deterioration analysis of flexible pavements under overweight vehicles. *J. Transp. Eng.*, 133(11), 625–633.

[5] Pais, J.C., Amorim, S.I.R., Minhoto, M.J.C. (2013). Impact of traffic overload on road pavement performance. *J. Transp. Eng.*, 873–879.

[6] Jacob, B., O'Brien, E., Jehaes, S. (2002). *COST 323: Weigh-in-Motion of road vehicles – final report*. Paris.

[7] Cebon, D. (200). *Handbook of vehicle-road interaction*. CRC Press, Taylor & Francis.

[8] Scheuter, F. (1998). Evaluation of Factors Affecting WIM System Accuracy. *COST 323 Weigh in Motion of Road vehicles, Proceedings of the Second European conference on WIM*.

[9] Gajda, J., Burnos, P., Sroka, R. (2018). Accuracy Assessment of Weigh-in-Motion Systems for Vehicle's Direct Enforcement. *IEEE Intell. Transp. Syst. Mag.*, 88–94.

[10] Burnos, P., Rys, D. (2017). The Effect of Flexible Pavement Mechanics on the Accuracy of Axle Load Sensors in Vehicle Weigh-in-Motion Systems. *Sensors*, 17(9), 2053.

[11] Gajda, J. Sroka, R. Stencel, J., Zeglen, T., Piwowar, P., Burnos, P. (2012). Analysis of the temperature influences on the metrological properties of polymer piezoelectric load sensors applied in Weigh-in-Motion systems. *2012 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 772–775.

[12] Burnos, P., Gajda, J. (2016). Thermal Property Analysis of Axle Load Sensors for Weighing Vehicles in Weigh-in-Motion System. *Sensors*, 16(12).

[13] Slavik, M. (2008). WIM accuracy verification through simulation. *Proceedings of the International Conference on Heavy Vehicles, 5th International Conference on Weigh-in-Motion of Heavy Vehicles*, 412–422.

[14] ASTM, E1318 – 09: Standard specification for highway Weigh-in-Motion (WIM) systems with user requirements and test methods, 2009.

[15] Koniditsiotis, C. (2000). *Weigh-in-Motion Technology*. Sydney.

[16] Wilks, S. (1941). Determination of Sample Sizes for Setting Tolerance Limits. *Ann. Math. Stat.*, 12(1), 91–96.

[17] Proschan, F. (1953). Confidence and Tolerance Intervals for the Normal Distribution. *J. Am. Stat. Assoc.*, 48(263), 550–564.

[18] *Measurement Specialties. 2014. http://www.meas-spec.com/*