

## Diagnosis of the Accuracy of the Vehicle Scale Using Neural Network

Tomáš Kliment<sup>1,2</sup>, Jaromír Markovič<sup>1</sup>, Dušan Šmigura<sup>1,3</sup>, Peter Adam<sup>1,3</sup>

<sup>1</sup> Slovak Legal Metrology, n.o., Hviezdoslavova, 31, 97401, Banská Bystrica, Slovakia, [kliment@slm.sk](mailto:kliment@slm.sk)

<sup>2</sup> Department of Aviation Technical Studies, Faculty of Aeronautics, Technical University of Košice, Rampová, 7, 04121, Košice, Slovakia

<sup>3</sup> Institute of automation, measurement and applied informatics, Faculty of Mechanical Engineering, Slovak Technical University, Námetie slobody, 17., 81231, Bratislava, Slovakia

The article describes a method for diagnosing the accuracy of the vehicle scale without using standard weights. The novel method defines the possibility to estimate whether the scale would pass the test for error of indication in the next verification or not, only by using the results from simple tests with load of estimated weight and appropriate classifier. The method is primarily developed for users of these scales. Created classifier is based on the neural network algorithm. The neural network was trained with data from verifications, which are provided by Slovak Legal Metrology. Well trained classifier can provide not only information whether the scale will potentially pass the mentioned test or not, but reliability which is associated with this result as well. In this way, the user has valuable information about the scale in the period between the verifications.

Keywords: The vehicle scale, accuracy, classifier, neural network, legal metrology.

### 1. INTRODUCTION

Activities connected with uniformity and correctness of measurements, methods and evaluation of measurement data are current and widely discussed topics. In general, metrology deals with these activities. Legal metrology deals with measurements that are connected with transparency of business transactions, consumer protection or environment protection. The purpose of the legal metrology institute is the protection of credibility of the measurement results in the government-controlled area. Significant part of the gross domestic product in developed countries is from the area of metrology and related areas [1].

Weighting of the vehicles (or vehicles with cargo) belongs to the government-controlled area, so vehicle scales are legally controlled measuring instruments [2]. Metrological aspects, calibration and evaluation of measurement data are described in [3]-[6]. Uncertainty evaluation is widely discussed in [7]-[14].

Validity period of verification of this measuring instrument is 2 years. During this period the scale is used many times and it is difficult to guarantee the same metrological characteristics as the scale had at the time of verification.

The need of information about the condition of the scale between verifications resulted in the development of the diagnostic method.

The diagnostic method is primarily designed for the user of the scale. For this reason, it should be simple and executable without standard weights, only with load of estimated weight.

Several tests are part of the verification of the vehicle scale. Tests include the eccentricity test and test for error of indication. Interesting finding is that 87.5 % of the scales that do not pass the eccentricity test do not pass the test for error of indication as well.

As a result, we can deduce an assumption that the eccentricity test is a good indicator whether the scale will pass the test for error of indication.

Although the mentioned tests are well-known, it is appropriate to briefly describe them. The eccentricity test is performed by the application of the adequate standard weight  $m_{\text{ref,ecc}}$  (mostly within the range 8000 – 12000 (kg)) on the defined eccentric point of the load receptor. Then we can define maximal error from the eccentricity test  $E_{\text{ecc,max}}$ :

$$E_{\text{ecc,max}} = \max \left\{ \left| I_j - m_{\text{ref,ecc}} \right| \right\} \quad (1)$$

$I_j$  is indicated value from the eccentric j-th point of the load receptor. It is analyzed whether the results are similar enough to the reference value of standard weight, regardless of the position of the test load. However, the user of the scale mostly

does not have the standard weights at disposal. Therefore, the test must be modified in a way that  $E_{ecc,max}$  is defined as a difference between maximum and minimum of the indicated value by weight of estimated value on the defined eccentric point of the load receptor.

$$E_{ecc,max} = \max(I_j) - \min(I_j) \quad (2)$$

The load receptor is uniformly loaded with predefined values from minimum to maximum and from maximum to minimum (with the same values of standard weights) during the test for error of indication. In this way, the error of indication of the scale is defined in the whole measuring range of the measuring instrument. The maximum of the error  $E_{ind,max}$  from the test for error of indication can be defined as:

$$E_{ind,max} = \max\{|I_i - m_{ref,i}|\} \quad (3)$$

Where  $I_i$  is indicated value of scale in  $i$ -th point and  $m_{ref,i}$  is standard weight in  $i$ -th point.

First, we tried to define the analytical formula between the results from the eccentricity test and the test for error of indication. More specifically, it was the formula between the values  $E_{ecc,max}$  and  $E_{ind,max}$ . In Fig.1. and Fig.2. is shown  $E_{ecc,max}$  vs.  $E_{ind,max}$ .

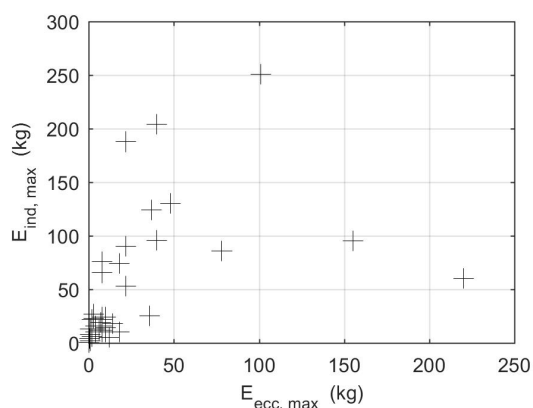


Fig.1. Correlation between  $E_{ecc,max}$  and  $E_{ind,max}$  (chosen data from verifications).

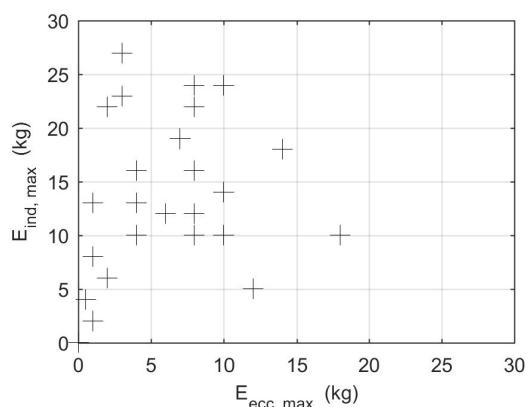


Fig.2. The detail of correlation.

The visualization of the data shows that there is no evident correlation. It is necessary to choose another approach. Based on this fact, the classifier, which will be able to classify the scale as accurate or inaccurate, with the available data, will be designed.

## 2. THEORY

Slovak Legal Metrology, n.o. is in the possession of a large database of records from verifications – about 1500 verifications of the vehicle scales (during a period of about last 2 years). Some of these data will be used as an input dataset for the classifier. The neural network [15]-[20] was chosen as the classifier. The neural network operates in two modes: learning and working one. In the learning mode, it learns on the dataset from the mentioned database, hence it is learning with known inputs and known outputs. The neural network will try to adjust a limit value for  $E_{ecc,max}$ , which will be able to divide scales in two classes, those which would potentially pass the test for error of indication (or we can say accurate scales) or those which would not (inaccurate scales). The  $E_{ecc,max}$  parameter was chosen for this purpose, because small value is a good indicator of the accurate scale. Subsequently, the learned neural network saves its parameters, called weights. Since neural network can work with known inputs and unknown outputs, the output of the network is the Bayesian a posteriori probability that new scales belong to the class of the accurate scales. The main task is how to optimally train neural network as the classifier for simple usage. The inputs to the neural network are:

1. the maximum error from the eccentricity test  $E_{ecc,max}$  defined in equation (2),
2. maximum capacity of the scale  $max$ ,
3. scale interval  $d$ ,
4. standard weight for eccentricity test  $m_{ref,ecc}$ .

The neural network is trained with  $m_{ref,ecc}$ . After learning process, user will use this method with  $I_{mid}$  (indicated value of weight of estimated value from the middle of the load receptor) instead of  $m_{ref,ecc}$ .

These data create the input column vector  $\mathbf{X}$ :

$$\mathbf{X} = \begin{bmatrix} E_{ecc,max} \\ max \\ d \\ m_{ref,ecc} \end{bmatrix} \quad (4)$$

The sequence of these vectors creates the dataset for the network. Input data are visualized in Fig.3.:

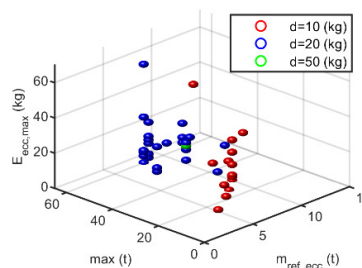


Fig.3. Input data of the neural network.

The used network is two-layered. It is composed of the input, hidden and output layer. The input layer performs data decorrelation. There is no universal rule to determine the optimal number of parameters for a network [15], but it is important to consider a fact, that the number of parameters of the whole neural network should be smaller than the number of inputs. In this case, the network has 25 parameters with 43 inputs. Inappropriate large network can cause overfitting [21], [22]. The result is: 4 units in the hidden layer with the sigmoidal activation function and 1 unit in the output layer. The output of the hidden layer is column vector  $\mathbf{O}_1$ :

$$\mathbf{O}_1 = s(\mathbf{W}_1 \mathbf{X} + \mathbf{B}_1) \quad (5)$$

Where matrix  $\mathbf{W}_1$  contains weights from hidden layer,  $\mathbf{B}_1$  is column vector of biases of hidden layer and  $s$  is a sigmoidal function.

$$\mathbf{W}_1 = \begin{bmatrix} \omega_{11} & \omega_{21} & \omega_{31} & \omega_{41} \\ \omega_{12} & \omega_{22} & \omega_{32} & \omega_{42} \\ \omega_{13} & \omega_{23} & \omega_{33} & \omega_{43} \\ \omega_{14} & \omega_{24} & \omega_{34} & \omega_{44} \end{bmatrix} \quad \mathbf{B}_1 = \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ b_{14} \end{bmatrix} \quad (6)$$

The output layer contains 1 unit with softmax activation function. Network output  $p$  (number from 0 to 1) is the Bayesian a posteriori probability whether the scale is accurate or not. This a posteriori probability is defined by:

$$p = \text{softmax}(\mathbf{O}_1^T \mathbf{W}_2 + \mathbf{B}_2) \quad (7)$$

$\mathbf{W}_2$  is column vector of the weights of the output layer and  $\mathbf{B}_2$  is bias of the output layer.

$$\mathbf{W}_2 = \begin{bmatrix} \omega_{21} \\ \omega_{22} \\ \omega_{23} \\ \omega_{24} \end{bmatrix} \quad \mathbf{B}_2 = [b_{21}] \quad (8)$$

We used dichotomous classification, when each output  $p$  which has greater value than 0.5 means accurate scale (belonging to class of the accurate scales, marked as 1) and value smaller than 0.5 means inaccurate scale (belonging to the complement of the accurate scales class, marked as 0) [23]. The scheme of the used neural network is shown in Fig.4.

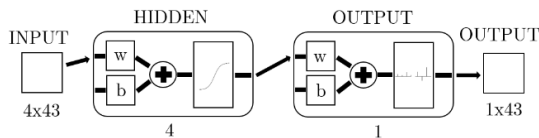


Fig.4. Classification neural network.

Neural network uses the scaled conjugate gradient propagation algorithm as a learning rule [24]. During the learning, the network tries to find optimal values of  $\mathbf{W}_1$ ,  $\mathbf{W}_2$ ,  $\mathbf{B}_1$ , and  $\mathbf{B}_2$ .

### 3. RESULTS

The input dataset contained 43 scales, whereas 16 were inaccurate and 27 accurate. This dataset is relatively small for sufficient training of the neural network. It is due to a limited number of the scales which do not pass the given tests and because the input space must be uniformly distributed (approximately the same number of accurate and inaccurate scales in the input dataset), that the training dataset is quite small.

The dataset was divided into 3 sub-datasets: training, validation, and test set.

Stop criterion was the point where mean squared error MSE on the validation dataset reached the minimum for the previous epochs and for the next 100 epochs. Epoch is a part of learning period, when network learns with whole data from the training dataset. In this way, we can avoid overtraining situation, a situation when neural network does not learn only with systematic components but with random components as well.

The results from the learning process are shown as a "confusion matrix" [25] for all datasets.

Table 1. "Confusion matrix" for training dataset.

Output class	0	8	1	88.9%
	1	1	21	95.5%
		88.9%	95.5%	93.5%
		0	1	
Target class				

Table 2. "Confusion matrix" for validation dataset.

Output class	0	1	0	100%
	1	2	3	60.0%
		33.3%	100%	66.7%
		0	1	
Target class				

Table 3. "Confusion matrix" for test dataset.

Output class	0	3	0	100%
	1	1	2	66.7%
		75.0%	100%	83.3%
		0	1	
Target class				

Table 4. "Confusion matrix" for all datasets.

Output class	0	12	1	92.3%
	1	4	26	86.7%
		75.0%	96.3%	88.4%
		0	1	
Target class				

The target class can be: 0 – the scale, which does not pass the test for error of indication or inaccurate scale, 1 – the scale which passes the test for error of indication or accurate scale. Output class can be 0, when a posteriori probability  $p$  is smaller than 0.5 or 1, when  $p$  is bigger than 0.5. For instance,  $p = 0.51$  means the scale is rather accurate and  $p = 0.49$  means the scale is rather inaccurate, but this classification is unreliable. In the ideal case, the target class is equal to output class for all data. In the green field is the number of the scales, which the network classified correctly. In the red field is the number of scales classified incorrectly.

Optimal network weights of hidden layer after learning process were adjusted as:

$$\mathbf{W}_1 = \begin{bmatrix} -1.086 & 0.610 & -0.667 & -1.205 \\ -0.010 & 1.535 & -1.181 & -0.699 \\ 5.651 & -0.641 & -0.111 & 0.427 \\ 2.614 & 2.180 & -0.282 & -0.202 \end{bmatrix} \quad \mathbf{B}_1 = \begin{bmatrix} 2.178 \\ -1.264 \\ 5.505 \\ 2.483 \end{bmatrix} \quad (9)$$

Optimal network weights of output layer:

$$\mathbf{W}_2 = \begin{bmatrix} 0.020 \\ 0.623 \\ -8.247 \\ -3.185 \end{bmatrix} \quad \mathbf{B}_2 = [0.480] \quad (10)$$

For better graphical visualization of the output function  $p$  (a posteriori probability) defined in the equation (7), we set inputs defined in equation (4)  $d$  and  $max$  as the parameters.  $E_{ecc,max}$  and  $m_{ref,ecc}$  were set as the variables. Fig.5. is for the case of vehicle scale with  $max = 30000$  (kg) and  $d = 10$  (kg). Fig.6. is for the scale of  $max = 60000$  (kg) and  $d = 20$  (kg).

As we can see, the output function is rippled. It is caused by the fact that the neural network is trained for values  $m_{ref,ecc} \in \{8000, 12000\}$  (kg), what is the range for the test load of eccentricity test. Outside this interval the output function is unreliable. It is generally known that the regression network should be used for such inputs as it has been previously trained. It results in the fact that the user should use this algorithm with the test load from 8000 (kg) to 12000 (kg). In that case, when the user does not have such test load at his disposal, he might linearly extrapolate the  $E_{ecc,max}$ . Only in that case  $E_{ecc,max}$  will be in the desired range.

Second test dataset with scales without service before last verification was created. These are the best data for testing the correctness of the algorithm, because they reflect the long-time period between verifications.

Table 5. “Confusion matrix” for second test dataset.

Output class	0	2	0	100%
	1	0	9	100%
		100%	100%	100%
		0	1	

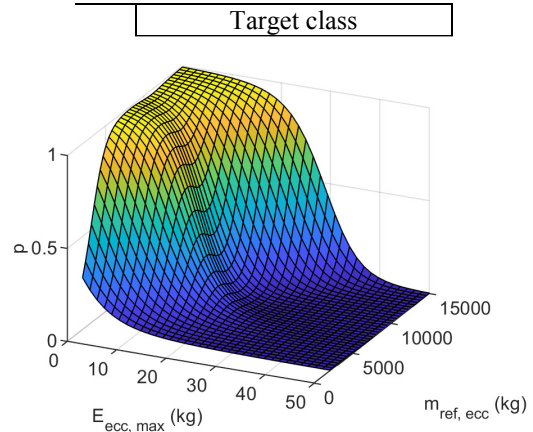


Fig.5. Output function for scale with  $max = 30000$  (kg) and  $d = 10$  (kg).

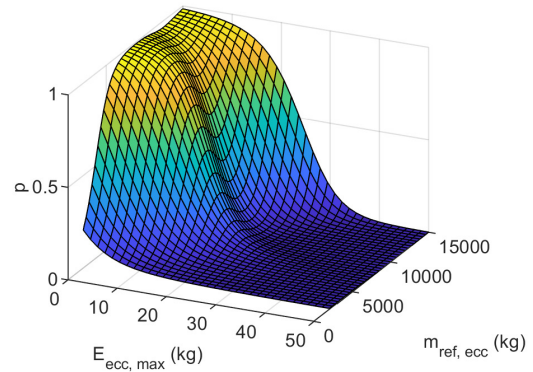


Fig.6. Output function for scale with  $max = 60000$  (kg) and  $d = 20$  (kg).

After the satisfying training and testing, we can use the neural network as the classifier. The use is shown in the next block diagram, Fig.7.:

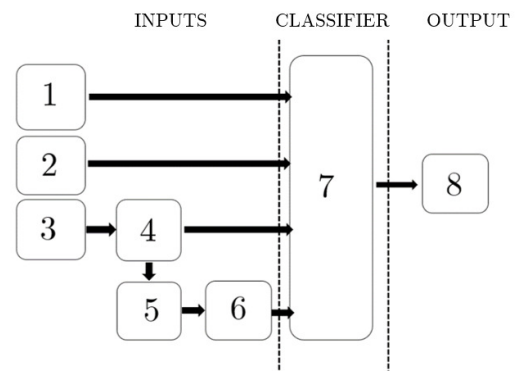


Fig.7. Block diagram of a posteriori probability determination.

- 1 – update of the classifier with new information
- 2 – information about the scale:  $max, d$
- 3 – load of the estimated weight
- 4 – indicated value in the middle of the load receptor  $I_{mid}$
- 5 – modified eccentricity test
- 6 – the calculation of the  $E_{ecc,max}$

7 – a posteriori probability determination by classifier

8 – a posteriori probability whether scale is accurate or not

The user performs only the eccentricity test in the modified form and with the maximum capacity max, scale interval  $d$  and indicated value  $I_{mid}$  has all the needed input information for the classifier. The classifier can be implemented on the online website, which is very user friendly. With the growing number of new verifications, it is possible to create new and larger datasets, so that the classifier can be updated and can improve its classification performance.

#### 4. CONCLUSIONS

The goal of this article was to verify the options of the diagnostics of the vehicle scale only with test load of estimated weight, which is primarily designed for the user of the scale. The best method for such verification is certainly with the standard weights, however, the user usually does not dispose of it.

With this method, the user can diagnose the condition of the scale during the period between verifications. It is important for the question of accuracy, because the owner of the scale protects himself from the loss of profit, protects the state to which he levies the taxes and protects his clients.

Data from the verifications were used to create the classifier. After the simple test with test load of estimated weight, the classifier can determine a posteriori probability, whether the scale is accurate or not, hence whether the scale would pass the test for error of indication or not in the next verification.

The experiment confirmed that the reliability of this classifier is 83.3 %, which is the rate of the correct classification on the testing dataset. Moreover, additional available test dataset created from scales without the service has 100 % reliability.

It is assumed that the accuracy of the classification will be improved by acquisition of more appropriate data from verifications and by the division of the scales according to the industry in which they are used. Different results could be provided by the classifier trained on the scales used, e.g. in stone quarry and by the classifier trained on the scales used in the stores. Specific operating conditions can impact mechanical parts of the vehicle scales [26]. These improvements should increase reliability of the classifier. The reliability 83.3 % for a dataset this small only confirms the quality of the proposed classifier.

Classification function is simple to implement. The user who wants to diagnose his vehicle scale needs to know only the scale interval, maximum capacity, indicated value in the middle of the scale and the result of the modified eccentricity test.

#### ACKNOWLEDGMENT

This work was supported by the APVV 15-0164.

#### REFERENCES

- [1] Birch, J. (2003). *Benefit of legal metrology for the economy and society*. A study for the International Committee of Legal Metrology.
- [2] International Organization of Legal Metrology. (2010). *Alphabetical list of terms defined in OIML Recommendations and Documents*. OIML G 18.
- [3] Zeier, M., Allal, D., Judaschke, R. (2011). *Guidelines on the evaluation of vector network analysers (VNA)*. EURAMET Calibration Guide No. 12. Braunschweig, Germany: EURAMET e.V.
- [4] EURAMET Technical Committee for Mass and Related Quantities. (2015). *Guidelines on the Calibration of Non-Automatic Weighing Instruments*. EURAMET Calibration Guide No. 18. Version 4.0 (11/2015). Braunschweig, Germany: EURAMET e.V.
- [5] European Cooperation in Legal Metrology. (2015). *Software guide. (Measuring instruments directive 2014/32/EU)*. WELMEC 7.2.
- [6] European Cooperation in Legal Metrology. (2015). *Directive 2009/23/EC: Common application. Non-automatic weighing instruments*. WELMEC 2, Iss. 6.
- [7] Joint Committee for Guides in Metrology (JCGM). (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement*. JCGM 100:2008.
- [8] Joint Committee for Guides in Metrology (JCGM). (2011). *Evaluation of measurement data - Supplement 2 to the "Guide to the expression of uncertainty in measurement" - Extension to any number of output quantities*. JCGM 102:2011.
- [9] Joint Committee for Guides in Metrology (JCGM). (2008). *Evaluation of measurement data - An introduction to the "Guide to the expression of uncertainty in measurement" and related documents*. JCGM 100:2008.
- [10] Joint Committee for Guides in Metrology (JCGM). (2012). *Evaluation of measurement data - The role of measurement uncertainty in conformity assessment*. JCGM 106:2012.
- [11] Witkovský, V., Wimmer, G., Ďuriš, S. (2015). On statistical methods and reference confidence intervals in interlaboratory comparisons for temperature. *International Journal of Thermophysics*, 36 (8), 2150-2171.
- [12] Palenčár, R., Sopkuliak, P., Palenčár, J., Ďuriš, S., Suroviak, E., Halaj, M. (2017). Application of Monte Carlo Method for evaluation of uncertainties of ITS-90 by Standard Platinum Resistance Thermometer. *Measurement Science Review*, 17 (3), 108-116.
- [13] Palenčár, R., Wimmer, G., Halaj, M. (2002). Determination of the uncertainties and covariances in the calibration of the set of weights. *Measurement Science Review*, 2 (1), 9-20.
- [14] Charvátová, H., Janáčková, D., Fialka, M., Kolomazník, K. (2010). Mass analysis of the components separated from printed circuit boards. *Acta Montanistica Slovaca*, 15 (1), 58-61.

- [15] Rojas, R. (1996). *Neural Networks: A Systematic Introduction*. Springer.
- [16] Haikyn, S. (2008). *Neural Networks and Learning Machines*, (3rd ed.). Pearson.
- [17] Sutkens, J.A.K., Vandewalle J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, 9, 293-300.
- [18] Huang, G., Song, S., Wu, C., You, K. (2012). Robust support vector regression for uncertain input and output data. *IEEE Transactions on Neural Networks and Learning Systems*, 23, 1690-1700.
- [19] Vapnik, V. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10, 988-999.
- [20] Nyulászai, L., Andoga, R., Butka, P., Gašpar, V. (2016). Comparison of experimental identification methods using measured data from a turbojet engine. In *IEEE 14th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*. IEEE, 23-37.
- [21] Tetko, I.V., Livingstone, D.J., Luik, A.I. (1995). Neural network studies. 1. Comparison of overfitting and overtraining. *Journal of Chemical Information and Computer Sciences*, 35, 826-833.
- [22] Chicco, D. (2017). Ten quick tips for machine learning in computational biology. *BioData Mining*, 10 (35).
- [23] Rojas, R. (1996). A short proof of the posterior probability property of classifier neural networks. *Neural Computation*, 8 (1), 41-43.
- [24] Møller, M.F. (1991). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6, 525-533.
- [25] Congalton, G.C. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37 (1), 35-46.
- [26] Dovica, M., Buša, J., Palenčár, R., Ďuriš, S., Šooš, E., Vrba, I., Kelemenová, T., Škovránek, T. (2013). Comparison of methods for analysis of deviations from roundness. *Measurement Techniques*, 56 (9), 1021-1025.

Received May 28, 2018

Accepted January 21, 2019