An Approach for the Evaluation of a Measurement System: A Study on the Use of Machine Learning and Predictions

Malinka Ivanova
Department of Informatics, Faculty of Applied Mathematics and Informatics, Technical University of Sofia, Bulgaria
m_ivanova@tu-sofia.bg

Valentin Tsenev
Department of Electronics and Energy Engineering, Technical College of Sofia, Technical University of Sofia, Bulgaria
vtsenev@tu-sofia.bg

Vania Mikova
Faculty of Engineering and Pedagogy – Sliven, Technical University of Sofia, Bulgaria
vanya_mikova@abv.bg

Received: 27 September 2023 | Revised: 16 October 2023 | Accepted: 20 October 2023

 DOI: https://doi.org/10.48084/etasr.6450

ABSTRACT

Quality control during the manufacturing process is an important factor in delivering products in electronics according to planned characteristics and properties. It concerns the capability of the chosen measurement system to perform precise and reliable measurement trials, which is evaluated mainly through the utilization of measurement system analysis. In order to reduce time effort and to partially automate these operations, a methodology for the prediction of a part of the dataset through applying the Neural Net algorithm is proposed in this paper in two scenarios: (1) when two metrology experts are involved in the measurement in three trials and the data of a third specialist are predicted and (2) when three metrology specialists collect data in two trials and the data of the third trial are predicted. The developed predictive models in these two scenarios are assessed and they are characterized by high accuracy. Gage repeatability and reproducibility analysis are used to evaluate the measurement systems based on original and partially artificial datasets as the comparative results outline the suitability of the proposed approach, due to the proximity of the obtained values.

Keywords-machine learning; semi artificial dataset; measurement system; measurements system analysis; Gage repeatability and reproducibility; electronics manufacturing; automation

I. INTRODUCTION

In electronics, quality control during the manufacturing of a certain product is of particular importance. It guarantees the delivery by the manufacturer and the use by the user of a product possessing the previously announced characteristics and properties [1, 2]. In this way, low-quality production is prevented from reaching the consumer and additional costs for the manufacturer are avoided. Different tools for improving manufacturing systems are discussed in [3], showing existing practices and some challenging issues. Quality control is related to the choice of a system for measuring product-typical parameters, which must meet certain criteria, such as precision, accuracy, stability, and reliability. Thus, measurement systems must be evaluated as this is conducted mainly through the utilization of statistical methods. The most popular statistical approach for the evaluation of a measurement system is Measurement System Analysis (MSA), whose purpose is to identify the dependence between variances at measurement and the variability of the measurement process as a whole. This statistical method is evolving to satisfy different requirements of the measuring process or the measured product, due to the development and integration of a wide variety of technologies in the manufacturing process. Authors in [4] propose MSA as a one-side tolerance method. Authors in [5] present a new procedure regarding the MSA methodology for real time evaluation and for identifying MSA suitability for quality control in a manufacturing process, allowing continuous monitoring and adjustment of the measurement system and in combination with Statistical Process Control (SPC), it is possible to achieve high-quality manufacturing. Authors in [6] discuss a holistic approach for applying the Gauge study of high-density data collected through the usage of a 3D laser scanner (so called point cloud). The investigation reveals an approach for analyzing the repeatability and reproducibility of point clouds and for uncertainty quantification of these points.

Well-known statistical methods for the evaluation of measurement systems are widespread in practice and discussed among the scientific community, which shows their effectiveness in producing quality products. We are looking for an approach that could reduce the time and effort of metrology experts. For this purpose, the current investigation is focused on whether one semi artificially generated measurement system possesses identical or similar characteristics with the original one. Thus, the aim of the current paper is to present a methodology for evaluating measurement systems with partially artificial datasets. Statistical technique Gage Repeatability and Reproducibility (GR&R) is applied for studying the measurement system and machine learning is used for predicting part of data in the measured datasets.

II. THE PROPOSED METHODOLOGY

A. *Description*

The proposed methodology is depicted in Figure 1. It consists of the following steps:

1. Measurement data are collected in original datasets by two or three operators A, B, and/or C (metrology specialists).
2. The original datasets for three different dimensions of the product are processed in Minitab software through applying GR&R analysis and the ANOVA method.
3. The obtained results after GR&R analysis are used for the evaluation of the measurement system.
4. Machine learning algorithm Neural Net in Rapid Miner Studio is applied to the original datasets to make predictions regarding measurement data of operator C or regarding the third trial of the operators.
5. Datasets with predictions are the base for GR&R analysis.
6. The achieved results after GR&R analysis of the partially artificial measurement system are evaluated.
7. The two measurement systems are compared based on the original and on partially artificial datasets.

The methodology is developed for two scenarios depending on which data EW artificially created by the predictions.

The first scenario is related to the case when the collected data are measured by operators A and B in three trials. The data of the third operator C in three trials are artificially generated through predictions based on the historical data of operators A and B. The originally collected data are colored in blue color in Figure 2 and the artificially predicted data in green. Each trial includes ten different measurements. In the second scenario, data are gathered from three operators A, B, and C in two trials and the data of the third trial of operators A, B, and C are artificially generated.

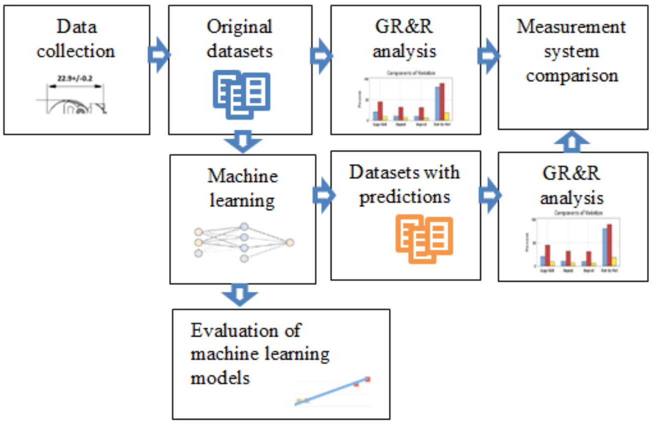

Fig. 1. The proposed methodology.

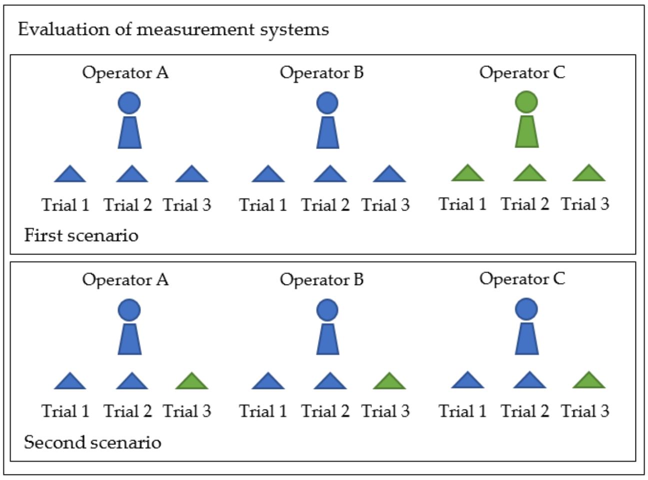

Fig. 2. The two considered scenarios in the proposed methodology.

B. *Controlled Object and Measurement Machine*

The controlled object is the thermo-printing mechanism SS205 V.4 used for heavy vehicle tachographs for the automotive industry. The measured dimensions of SS205 V.4 used in this study are presented in Figure 3 and they are: first dimension: 25.5 mm, second dimension: 22.9 mm, and third dimension: 67.7 mm with acceptable variance of +/-0.2 mm. The measurement system comprises 3D coordinate measuring machine Hexagon DEA Global Silver Performance 07.07.05 which runs PC-DMIS software version 2018 R1.

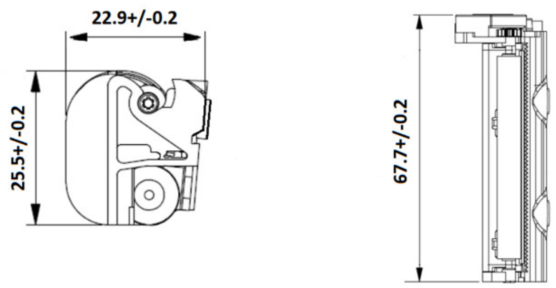

Fig. 3. Measured dimensions of the thermo-printing mechanism SS205 V.4

C. Data Collection

The data used for the analysis are 10 samples of the final product SS205-V4 concerning 3 different dimensional parameters, measured by 2 metrology specialists 3 times. Some of the measured data are shown in Table I.

TABLE I. MEASURED DATA BY TWO METROLOGY SPECIALISTS IN THREE TRIALS

Part	Trial	Operator	25.5 mm First dataset	22.9 mm Second dataset	67.7 mm Third dataset
1	1	A	25.378	22.890	67.672
1	2	A	25.379	22.875	67.673
1	3	A	25.378	22.879	67.682
2	1	A	25.386	22.891	67.663
2	2	A	25.387	22.883	67.660
…	…	…	…	…	…

The data used in the second scenario for the analysis concern 10 samples of the final product SS205-V4 taking into account 3 different dimensional parameters, measured by 3 metrology specialists in 2 trials. Some of the measured data are shown in Table II.

TABLE II. MEASURED DATA BY THREE OPERATORS IN TWO TRIALS

Part	Trial	Operator	25.5 mm First dataset	22.9 mm Second dataset	67.7 mm Third dataset
1	1	A	25.378	22.890	67.672
1	2	A	25.379	22.875	67.673
2	1	A	25.386	22.891	67.663
2	2	A	25.387	22.883	67.660
3	1	A	25.355	22.879	67.669
	…	…	…	…	…

D. Applying Machine Learning

1) First Scenario

When evaluating a measurement system, measurements made by more than two experts are usually required. We have data from two metrology specialists and therefore the data from the third expert are artificially created by applying the machine learning algorithm Neural Net. Training and result obtaining are performed in RapidMiner Studio [7]. The prediction accuracy is very high as it is evaluated through standard machine learning metrics: Root Mean Squared Error (RMSE), Absolute Error (AE), Relative Error (RE), and Root Relative Squared Error (RRSE). The results can be seen in Table III.

TABLE III. PREDICTION ACCURACY OF MACHINE LEARNING MODELS BASED ON DATA OF TWO OPERATORS IN THREE TRIALS

Accuracy parameter	Prediction for 1st dimension	Prediction for 2nd dimension	Prediction for 3rd dimension
RMSE	0.0027	0.0043	0.0285
AE	0.0023	0.0035	0.0205
RE	0.0001	0.0002	0.0003
RRSE	0.5522	0.4319	0.8448

2) Second Scenario

The data for the third trial of the three operators are predicted through the Neural Net algorithm. The prediction accuracy is also very high.

III. EXPERIMENTATION AND RESULTS

A. Evaluation of the Measurement System in the First Scenario

The measured data are collected by operators A and B in 3 trials for each operator. The data for the third operator C are artificially generated. The measurement system evaluation is done through the usage of Minitab, applying GR&R analysis and the ANOVA method and the results regarding the first (25.5 mm), second (22.9 mm) and third (67.7 mm) dimensions are presented respectively in Figures 4, 5, and 6, respectively. Variance components and Gage evaluation metrics for the first dimension of the product are shown in Tables IV and V. The main used indicators are the %Contribution of variance components and %Study variations. %Contribution is applied to assess the variations regarding the measurement error of each source and %Study variations present a comparison regarding the variations of the measurement system and total variations.

For the first dimension (25.5 mm), the biggest variance is received for repeatability 0.68% as there is no variance for reproducibility and operator. Considering the measurements of the second (22.9 mm) and third (67.7 mm) dimensions, all sources present variations as the values are bigger for the second dimension (second dimension: repeatability 10.33%, reproducibility 9.91%, and operator 9.91%. Third dimension: repeatability 1.63%, reproducibility 1.84%, and operator 0.08%).

Repeatability concerns the variability of measurement when the same operator measures a given part several times. Repeatability is very small for the first dimension (0.68%), small for the third dimension (1.63%), and bigger for the second dimension (10.33%).

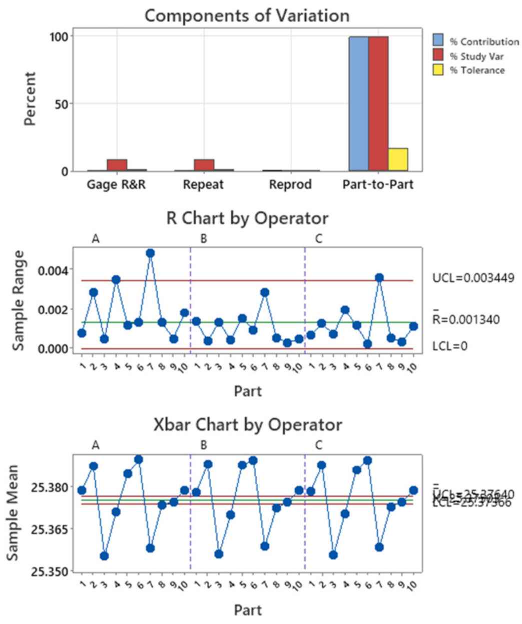

Fig. 4. Evaluation of the measurement system for the first dimension.

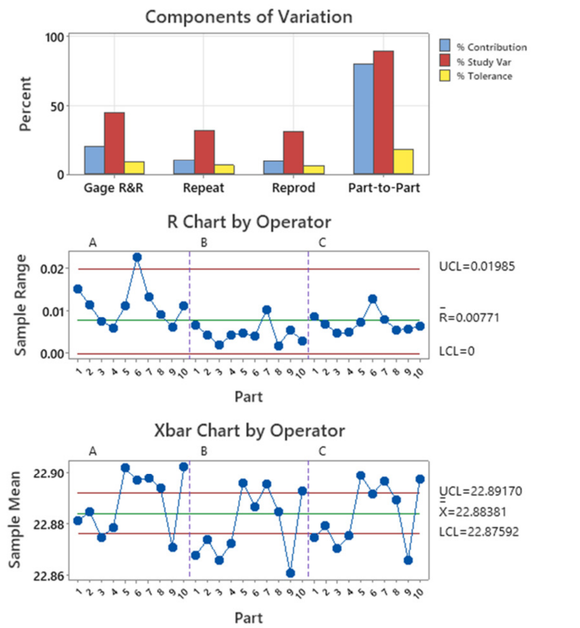

Fig. 5. Evaluation of the measurement system for the second dimension.

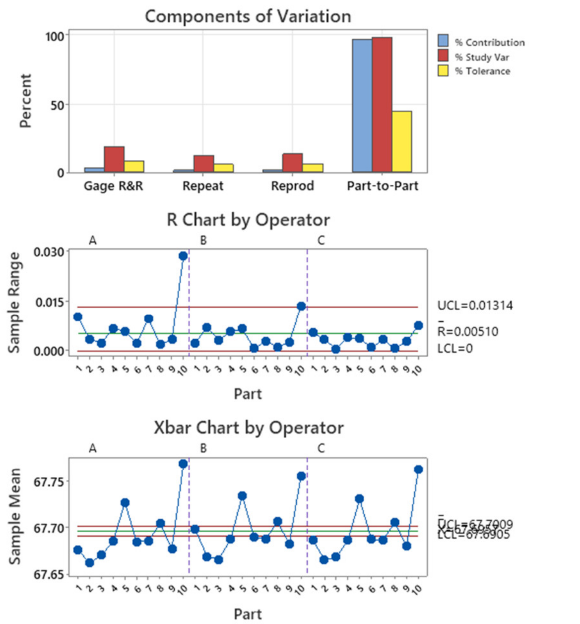

Fig. 6. Evaluation of the measurement system for the third dimension.

Reproducibility introduces measurement variability when several operators measure a given part. The value of reproducibility is zero for the first dimension, small for the third dimension (1.84%), and bigger for the second dimension (9.91%). When different parts have to be measured and there is variability in the measurements, then the variation source is part-to-part. The practice shows that the biggest variance in a measurement system introduced by the part-to-part component for different dimensions are: first dimension 99.32%, second dimension 79.76%, and third dimension 96.53%.

TABLE IV. VARIANCE COMPONENTS METRICS FOR THE FIRST DIMENSION 25.5 mm

Variance Components		
Source	**VarComp**	**%Contribution (of VarComp)**
Total Gage R&R	0.00000009	0.68
Repeatability	0.00000009	0.68
Reproducibility	0.00000000	0.00
Operator	0.00000000	0.00
Part-to-part	0.0001335	99.32
Total variation	0.0001345	100.00

TABLE V. GAGE EVALUATION FOR THE FIRST DIMENSION 25.5 mm

Gage Evaluation				
Source	**StdDev (SD)**	**Study Var (6xSD)**	**%Study Var (%SV)**	**%Tolerance (SV/Toler)**
Total Gage R&R	0.0009595	0.0057570	8.27	1.44
Repeatability	0.0009595	0.0057570	8.27	1.44
Reproducibility	0.0000000	0.0000000	0.00	0.00
Operator	0.0000000	0.0000000	0.00	0.00
Part-to-part	0.0115557	0.0693339	99.66	17.33
Total variation	0.0115954	0.0695725	100	17.33

The R chart indicates the operator consistency during measurements. It is seen that operator A is characterized with inconsistency at measurement of the first, second, and third dimension, because there are points above the Upper Control Limit (UCL) line and operators B and C measure parts consistently. The Xbar chart shows the part-to-part variations and concerns the repeatability component as the lines of average values and UCL and LCL (Low Control Limit) are drawn. Measurement by part chart presents variations when several parts are measured as the measurement system is better when the dots are closer for each part. The worst situation here is for the second dimension. Figure 7 presents the metrics of the first dimension.

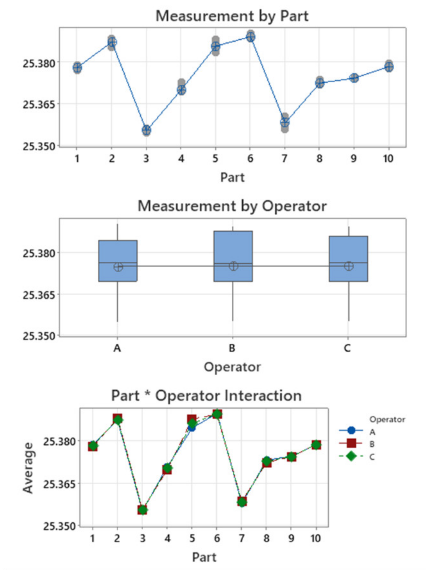

Fig. 7. Measurement by part, measurement by operator, and part*operator interaction charts for the first dimension (25.5 mm).

The chart measurement by operator in the ideal case must be a horizontal straight line that outlines similarity at the measurement of operators. The measurement of the second dimension by the three operators is characterized with bigger difference by operators in comparison to the measurements of the first and third dimensions. The chart part*operator interaction chart presents the average measurement for the operators A, B, and C and whether they measure the parts in a similar way. Similarity is almost achieved for the measurement of the first and third dimension, while in the second dimension there is a tolerance at part measurement. It can be said that the operators A, B, and C make the measurements in the same way, whereas for the second dimension, bias is added.

The total GR&R at variance component is 0.68% for the first dimension, 20.24% for the second, and 3.47% for the third dimension. The total GR&R evaluation regarding the percentage variation is 8.27% for the first dimension, 44.98% for the second, and 18.64% for the third dimension.

The acceptance of the measurement system is evaluated by taking into account the AIAG (Automotive Industry Activation Group) standards, which mention that the percentage contribution of variance components must be less than 1% for acceptance. If 1% < %Contribution < 9%, the measurement system is acceptable, but in the context of different factors and conditions. When %Contribution > 9%, the measurement system is unacceptable. So, from the results, it is seen that the measurement of the first dimension is acceptable, the third dimension is acceptable under conditions, and the second dimension measurement is unacceptable. Percentage study variations (%StudyVar) at Gage evaluation must be less than 10% for the measurement system to be acceptable, for acceptance under conditions the condition is 10% < %StudyVar < 30%, and the system is unacceptable at %StudyVar > 30%. In this way, the measurement system is acceptable for the first dimension, acceptable under conditions for the third dimension, and unacceptable for the second dimension (Table VI).

TABLE VI. MEASUREMENT EVALUATION OF THE THREE DIMENSIONS

Dimension	%Contribution	Evaluation	%StudyVar	Evaluation	Overall evaluation
First	0.68%	Acceptable	8.27%	Acceptable under conditions	Acceptable under conditions
Second	20.24%	Unacceptable	44.98%	Unacceptable	Unacceptable
Third	3.47%	Acceptable under conditions	18.64%	Acceptable under conditions	Acceptable under conditions

A comparison between the data taken by two operators and these by three operators is also made (the data of the third operator are artificially generated). The comparison is presented in Table VII. There is a slight decrease in %Contribution and %Study Variations values for three operator data in comparison to two operator data, but at both two and three operator datasets, the first and third dimensions are within the acceptable borders under conditions and the second dimension significantly exceeds the acceptable values and needs to be improved. It is seen from Table VII that the results for %Contribution at component variance and %Study variations at Gage evaluation for two and three operators are very similar and they could be used for decision making with similar success.

TABLE VII. COMPARISON THE RESULTS AT TWO AND THREE OPERATORS

Dimension	GR&R analysis	3 operators	2 operators
First	%Contribution of variance components	0.68%	1.03%
	%Study variation of Gage evaluation	8.27%	10.14%
Second	%Contribution of variance components	20.24%	29.46%
	%Study variation of Gage evaluation	44.98%	54.28%
Third	%Contribution of variance components	3.47%	6.18%
	%Study variation of Gage evaluation	18.64%	24.87%

B. *Evaluation of the Measurement System in the Second Scenario*

Data were collected by operators A, B, and C in two trials. The third trial for each operator was artificially generated. In this scenario, all components give their contribution to the variances in measurement as repeatability and reproducibility are smaller for the first and third dimension, i.e. 0.50% and 0.18% for the first dimension and 0.70% and 4.48% for the third dimension. For the second dimension, the values are bigger: 6% and 18.42%. According to Total GR&R at variance components (%Contribution of VarComp), the measurement system is acceptable for the first dimension 0.68% < 1%, unacceptable for the second dimension 24.42% > 9%, and acceptable under conditions for the third dimension 5.18% < 9%. Considering the total GR&R at Gage evaluation (%StudyVar), the measurement system is acceptable for the first dimension 8.24% < 10%, unacceptable for the second dimension 49.41% > 30%, and acceptable under conditions for the third dimension 22.75%<30%.

The R chart shows that there is inconsistency in the measurements of operator A at the second dimension and of operator B at the third dimension. The chart measure by part presents variations at measurement of the ten parts, which are bigger for the second dimension. Again, the difference at the measurement of the second dimension by operators A, B, and C is characterized with no straight line on the chart measurement by operator. Different operators measure this dimension in different ways.

From the chart part*operator interaction, it is seen that the average measurement for the operators A, B, and C for the first and third dimension, similarity is almost achieved, while in the second dimension there is a variance at part measurement.

A comparison regarding the measurement of the first, second, and third dimensions in the second scenario is presented in Table VIII. Considering the AIAG standard, it can be said that the measurement system is acceptable, acceptable under conditions, and unacceptable for the first, third, and second dimension, respectively.

TABLE VIII. MEASUREMENT COMPARISON

Dimension	**% Contribution**	**Evaluation**	**%StudyVar**	**Evaluation**	**Overall evaluation**
First	0.68%	Acceptable	8.24%	Acceptable	Acceptable
Second	24.42%	Unacceptable	49.41%	Unacceptable	Unacceptable
Third	5.18%	Acceptable under conditions	22.75%	Acceptable under conditions	Acceptable under conditions

The results after GR&R analysis applied on the datasets with 2 and 3 trials (the third trial is predicted) of 3 operators are compared in Table IX. Slightly increased values are observed in the case of 2 trials. The similarity in results outlines the suitability for usage of the artificially generated datasets.

TABLE IX. RESULT COMPARISON FOR THREE OPERATORS AND TWO AND THREE TRIALS

Dimension	**GR&R analysis**	**3 trials**	**2 trials**
First	%Contribution of variance components	0.68%	0.96%
	%Study variation of Gage evaluation	8.24%	9.78%
Second	%Contribution of variance components	24.42%	26.97%
	%Study variation of Gage evaluation	49.41%	51.93%
Third	%Contribution of variance components	5.18%	5.40%
	%Study variation of Gage evaluation	22.75%	23.23%

IV. CONCLUSIONS

In this paper, a new methodology for the evaluation of measurement systems is presented, which is based on partially generated artificial data. Two scenarios are considered: (1) when data by two operators are collected in three trials and the data of the third operator are predicted and (2) when data by three operators are gathered in two trials and the third trial of each operator is predicted. It is proven that this methodology is applicable and reduces time and effort for measurement performed by metrology experts who have to conduct many repetitive operations with high attention. Also, such an approach allows the analysis and evaluation of measurement systems to be partially automated.

The created machine learning models through usage of Neural Net algorithm are characterized with high accuracy. The comparison between measurement systems that utilize original data and partially artificial data is done and the results outline the suitability of this approach regarding missing data and facilitating the metrology specialists in measurement systems analysis.

ACKNOWLEDGMENT

This research is supported by the Bulgarian National Science Fund in the scope of the project "Exploration of the application of statistics and machine learning in electronics" under contract number КП-06-Н42/1.

REFERENCES

[1] I. A. Memon, Q. B. Jamali, A. S. Jamali, M. K. Abbasi, N. A. Jamali, and Z. H. Jamali, "Defect Reduction with the Use of Seven Quality Control Tools for Productivity Improvement at an Automobile Company," *Engineering, Technology & Applied Science Research*, vol. 9, no. 2, pp. 4044–4047, Apr. 2019, https://doi.org/10.48084/etasr.2634.

[2] A. B. E. Aichouni, H. Abdullah, and F. Ramlie, "A Scientific Approach of using the DMAIC Methodology to Investigate the Effect of Cutting Tool Life on Product Quality and Process Economics: A Case Study of a Saudi Manufacturing Plant," *Engineering, Technology & Applied Science Research*, vol. 11, no. 1, pp. 6799–6805, Feb. 2021, https://doi.org/10.48084/etasr.4008.

[3] M. Alquraish, "Modeling and Simulation of Manufacturing Processes and Systems: Overview of Tools, Challenges, and Future Opportunities," *Engineering, Technology & Applied Science Research*, vol. 12, no. 6, pp. 9779–9786, Dec. 2022, https://doi.org/10.48084/etasr.5376.

[4] S. Kamil and B. Pawel, "Measurement system analysis for one-sided tolerance," *ITM Web of Conferences*, vol. 15, 2017, Art. no. 05011, https://doi.org/10.1051/itmconf/20171505011.

[5] M. Diering, A. Hamrol, and A. Kujawińska, "Measurement System Analysis Combined with Shewhart's Approach," *Key Engineering Materials*, vol. 637, pp. 7–11, 2015, https://doi.org/10.4028/www.scientific.net/KEM.637.7.

[6] R. Dastoorian and L. Wells, "Gauge capability studies for high-density data: SPC Phase 0," *Procedia Manufacturing*, vol. 48, pp. 105–113, Jan. 2020, https://doi.org/10.1016/j.promfg.2020.05.026.

[7] *RapidMiner Studio Manual*. RapidMiner, 2014.