

APPLICATION OF REPEATABILITY AND REPRODUCIBILITY STUDIES FOR TESTING THE CAPABILITIES OF COORDINATE MEASURING MACHINES

Miloš Ranisavljev¹⁾, Branko Štrbac^{✉1)}, Predrag Janković²⁾, Branislav Dudić³⁾,
Sara Havrlišan⁴⁾, Miodrag Hadžistević¹⁾

1) University of Novi Sad, Faculty of Technical Sciences, Department of Production Engineering,
Trg Dositeja Obradovića 6, Novi Sad, Serbia (✉ strbacb@uns.ac.rs)

2) University of Niš, Faculty of Mechanical Engineering, ul. Aleksandra Medvedeva 14, Niš, Serbia

3) Faculty of Management, Comenius University Bratislava, 820 05 Bratislava, Slovakia

4) University of Slavonski Brod, Mechanical Engineering Faculty, ul. 108. brigade ZNG 36, Slavonski Brod, Croatia

Abstract

When it comes to measurement systems applicable for high demand output, reliable and accurate measurements, coordinate measuring machines are the go-to measurement equipment in today's manufacturing environments. In production conditions, providing data on the expanded measurement uncertainty for each measured quality characteristic is a demanding task, and often more than one CMM is used for measuring the quality characteristics on the same workpieces. In order to evaluate the suitability of two coordinate measuring machines for measuring three quality characteristics on one product, an expanded gage repeatability and reproducibility study was carried out. The setup and analysis of the experimental design were performed in accordance with the AS13003_1:2015 series of standards. The conducted study aimed to determine the suitability of the selected measurement systems for the defined task and the amount of their variation within the total variation of the production process. Also, considering that repeated measurements on CMMs are performed automatically, without operator influence, unsatisfactory GR&R results may indicate problems related to the choice of measurement strategy.

Keywords: expanded gage R&R study, design of experiment, CMM.

1. Introduction

The demand for reliable and precise measurements continues to grow, with the automotive and aerospace sectors imposing the most stringent requirements on measurement systems. Variability in measurement results across the supply chain remains a persistent and widely discussed challenge [1, 2]. Advancements in production technologies and automation are driving a steady increase in the simultaneous use of *coordinate measuring machines* (CMMs) for inspecting mass-produced components [3].

Coordinate measuring machines represent a fundamental technology for dimensional quality inspection. Touch-trigger CMMs, in particular, are distinguished by their high precision, making them indispensable for quality control in the production of precision components [4]. Workpieces in the mentioned industries often have strict requirements in the form of narrow tolerances, and in such cases, in order to reliably determine compliance with the specification, it is necessary to express the measurement uncertainty along with the measurement result. Since in practice it is difficult and economically inefficient to calculate the measurement uncertainty for each individual quality characteristic, the information available to the end user, outside certified metrological laboratories, concerns the *maximum permissible errors* (MPE). However, the MPE is not a sufficient indicator of quality of measurement results.

A very useful method for assessing measurement quality is *measurement system analysis* (MSA). The purpose of MSA is to identify the total variation present in a measurement system and the contribution of that variation to the overall variation of the production process. The most commonly used study within MSA is the gage repeatability and reproducibility study [5].

Planning a gage repeatability and reproducibility study for a coordinate measuring machine is not a minor undertaking. The basic prerequisite is that the measuring equipment included in the study is calibrated and traceable to the relevant national or international standard. A number of influential factors on the measurement result must be taken into account. Considering the complexity of procedures for calculating measurement uncertainty (GUM, ISO 15530 series, etc.), for the evaluation of the suitability of the measurement system used for measuring quality characteristics on critical workpieces in the aerospace industry, the AS13003_1:2015 standard recommends conducting a gage repeatability and reproducibility study, with minimum requirements that measurement systems must meet depending on the criticality of the characteristic being verified. The study recommends the inclusion of environmental factors (temperature, humidity, contamination, vibrations. . .), location factors (different buildings, workstations. . .), personnel factors (different shifts, time of day, levels of experience), production process factors (fixtures, probes, additional equipment. . .), as well as the selection of workpieces that represent the variation of the production process. To increase productivity, more than one coordinate measuring machine is often used for a series of workpieces. In that case, AS13003_1:2015 recommends conducting an expanded gage repeatability and reproducibility study, in which the following factors are considered: environmental factors – variation in ambient temperature; measuring machines – variation in the measurement results of an individual machine and between machines; and production process variation through a representative sample of workpieces [6].

Minix *et al.* used a gage repeatability and reproducibility study for a comparative analysis of measurement results of three identical workpieces, performed by three operators, on two coordinate measuring machines. One of the machines was located in a university laboratory, while the other was in a production company laboratory. Four potential influencing factors were selected: type of quality characteristic, method of point sampling (point-by-point and scanning mode), type of programming (teaching without a *computer-aided design* (CAD) model and with a CAD model), and operator factor. The conducted analysis identified significant differences between two operators when measuring different features; statistical significance was also observed between the use of the machine in manual and automatic mode, while variations in the measurement results of workpieces with and without a reference model were minimal. Certain differences in the total variation of the observed measuring machines were found, but none exceeded 10%, which is the *Automotive Industry Action Group* (AIAG) recommendation [7].

Dev *et al.* conducted a gage repeatability and reproducibility study with the aim of determining the suitability of a coordinate measuring machine for measuring a quality characteristic on a single workpiece in both manual and automatic measurement modes. Four trainees participated in the

study. The objective of the study was to demonstrate the importance of automation in coordinate measuring machines, through significantly higher repeatability of results in automatic mode compared to manual mode. After analysing the results, it was found that by using the automatic measurement mode, the reduction in the total variation of the gage amounted to 55%, while repeatability improved by 89% [8].

Shirodkar and Rane demonstrated the implementation of a *gage repeatability and reproducibility* (GR&R) study on a coordinate measuring machine for measuring four different quality characteristics: length, external and internal diameter, as well as one geometric characteristic, parallelism, on 10 identical workpieces. The influencing factors they considered were the position of the workpieces within the measuring volume of the CMM, at three levels; three different probe tip diameters; and three different scanning speeds. They concluded that the overall contribution of reproducibility was small compared to the total contribution of measurement system variation, for the factors of location and probe tip scanning speed. The probe tip diameter also had an influence on the percentage contribution of reproducibility. The conclusion was that the applied measurement system may or may not be suitable for measuring quality characteristics depending on their criticality, *i.e.*, functionality [9].

Although the GR&R methodology is well-established in practice, it is not without limitations. Andrés Carrión García and Angela Grisales del Río critically examined the conventional approach to measurement system analysis, which treats the total contribution percentage of the measurement system % GR&R and the number of distinct categories, *number of distinct categories* (NDC) as separate criteria for assessing system suitability. The analysis presented in their study demonstrates that these two indicators are fundamentally correlated, making their use as independent parameters redundant. Since both % GR&R and NDC reflect the same underlying variability inherent to the measurement system, relying on both criteria simultaneously may lead to confusion in decision-making. By evaluating the current acceptance thresholds recommended by the AIAG, the study proposes that a unified criterion should simplify and clarify the assessment of measurement system adequacy [10].

This study presents an extended repeatability and reproducibility analysis of two coordinate measuring machines, aimed at determining whether the selected machines are suitable for evaluating three quality characteristics on a single product. The experimental design and its analysis seek to establish whether the measurement results obtained from the observed machines differ statistically. If the selected measurement systems are found to be inadequate - and considering that the measurement program is identical on both machines and executed automatically, without operator influence - it becomes necessary to verify the validity of the measurement strategy and implement corrective actions. In addition, the experimental design will reveal internal variations within individual machines (run-to-run machine variation, representing the repeatability component), variations between the two CMMs (CMM-to-CMM, representing the reproducibility component), as well as variations caused by changes in ambient temperature conditions.

2. Methodology

The aerospace industry relies heavily on measurement and inspection procedures to ensure that delivered workpieces and assemblies meet the required level of accuracy. As a highly demanding sector, it imposes strict requirements throughout the supply chain. In response to the need for harmonization, standards such as AS13003:2015 have been developed. Issued by the *Society of Automotive Engineers* (SAE), this standard defines the minimum requirements for measurement system analysis within the aerospace supply chain [6].

2.1. MSA requirements from AS1300_1:2015

The purpose of the AS13003_1:2015 standard is to define the application of measurement system analysis tools and acceptance criteria within the supply chain of aircraft engine manufacturers. It also provides guidance for effective implementation of variation reduction strategies in cases where measurement systems are deemed unacceptable. This standard outlines the minimum requirements for conducting measurement system analyses that record results for both continuous and attribute variables, in accordance with the product's technical specification. When considering specification requirements, the standard introduces a classification based on the importance of the quality characteristic under study. Table 1 presents the minimum requirements for a given measurement system relative to the functional significance of the quality characteristic on the workpiece.

Table 1. Minimum requirements for MSA.

Method	Feature Category		
	Critical	Major	Minor
Resolution	$\leq 10\%$ of total tolerance ³		
Accuracy ratio ²	Requirement = 10:1		Requirement = 4:1
Accuracy Error / Bias	$\leq 10\%$ of total tolerance		
Repeatability	$\leq 10\%$ of total tolerance	$\leq 20\%$ of total tolerance	$\leq 30\%$ of total tolerance ¹
Gage R&R	$\leq 10\%$ of total tolerance	$\leq 20\%$ of total tolerance	$\leq 30\%$ of total tolerance ¹
Computer driven measurement systems correlation	$\leq 10\%$ of total tolerance		$\leq 20\%$ of total Tolerance
Linearity ²	$<1\%$ of total tolerance		
¹ Repeatability and Gage R&R studies are only required for minor features when there has been an escape related to that feature or when required by the purchaser. ² Only required when specified by the purchaser or where the feature is a datum. ³ May permit lower resolutions provided an acceptable GR&R is obtained.			

2.2. Mathematical principles for gage R&R expanded study

The extended *repeatability and reproducibility* (R&R) study was designed to explain the variations in results provided by a given measurement system when subjected to external sources of variation [5]. The study divides the total variation of the dependent variable/response into variation attributable to the measurement system and variation due to the workpieces (1). Measurement system variation is calculated as the sum of the variances associated with the system's repeatability and reproducibility (2) [11]:

$$\sigma^2_{\text{Total variation}} = \sigma^2_{\text{Total Gage R\&R}} + \sigma^2_{\text{Part-to-part}}, \quad (1)$$

$$\sigma^2_{\text{Total Gage R\&R}} = \sigma^2_{\text{Repeatability}} + \sigma^2_{\text{Reproducibility}}. \quad (2)$$

Measurement result variation can be attributed to repeatability when the measurement process is repeated for the same quality characteristic using the same instrument while keeping all parameters constant. In the case of CMMs, the operator has minimal influence on the overall measurement variation, as the measurement program is executed automatically. Consequently, any observed

variation can be attributed either to the workpieces or to the machine’s inherent repeatability. This statement is substantiated by establishing and examining the relationship between selected factors and the dependent variable Y , modelled using a general linear model:

$$Y_{ijkl} = \mu + P_i + O_j + A_k + PO_{ij} + PA_{jk} + OA_{jk} + POA_{ijk} + \varepsilon_{ijkl}, \quad (3)$$

where Y_{ijkl} represents the measured value (response) for the i -th workpiece, the j -th operator, and the k -th additional factor, in the l -th repetition. The symbol μ denotes the overall mean value of all observations. P_i, O_j, A_k represent the effects of the workpiece factor, the operator, and the additional factor A , respectively, while $PO_{ij}, PA_{jk}, OA_{jk}, POA_{ijk}$ denote their interactions. Finally, ε_{ijkl} represents the random error, *i.e.*, the variation arising from noise in the measurement results that is not explained by any factor.

An essential component of the subsequent analysis is the *ANalysis Of VAriance* (ANOVA). The advantage of the ANOVA method, compared to the average and range method for calculating variability in repeatability and reproducibility studies, lies in its ability to account for interactions between factors. Table 2 presents the general framework of the ANOVA method for an extended repeatability and reproducibility study of a measuring instrument [12].

Table 2. General framework of the ANOVA method.

Source	DF	SS	MS	F Statistic
Part (P)	$p - 1$	SS_P	MS_P	$\frac{MS_P}{MS_{PO}}$
Operator (O)	$o - 1$	SS_O	MS_O	$d \frac{MS_O}{MS_{OA}}$
Additional factor (a)	$a - 1$	SS_A	MS_A	$d \frac{MS_A}{MS_{OA}}$
$P \times O$	$(p - 1) \cdot (o - 1)$	SS_{PO}	MS_{PO}	$\frac{MS_{PO}}{MS_{POA}}$
$P \times A$	$(p - 1) \cdot (a - 1)$	SS_{PA}	MS_{PA}	$\frac{MS_{PA}}{MS_{POA}}$
$O \times A$	$(o - 1) \cdot (a - 1)$	SS_{OA}	MS_{OA}	$\frac{MS_{OA}}{MS_{POA}}$
$P \times O \times A$	$(p - 1) \cdot (o - 1) \cdot (a - 1)$	SS_{POA}	MS_{POA}	$\frac{MS_{POA}}{MS_E}$
Repeatability (Error)	$p \cdot o \cdot a \cdot (r - 1)$	SS_E	MS_E	–
Total	$N - 1$	SS_T	–	–

The component p denotes the number of workpieces, o represents the number of operators, a refers to the levels of the additional factor A , r is the number of repetitions per “workpiece–operator–additional factor A ” combination, and N is the total number of observations.

2.3. Experimental set-up

2.3.1. Workpiece

The workpiece used in the study was a sleeve assembly, with a strict requirement for the coaxiality of the individual parts within the assembly, shown in Fig. 1b. The production process used for the assembly of parts was press fitting. Coaxiality of the sleeve (1) with an internal nominal diameter of 20.6 mm and asymmetrical tolerance was evaluated against the piston guide (2), with a cylindrical feature as base A. In addition to coaxiality, the quality characteristics measured on the

workpieces were the inner diameter of the sleeve, 20.6 mm, and the combined length of 74.65 mm of the two parts. The clamped workpiece, positioned inside the working volume of the coordinate measuring machine, is portrayed in Fig. 1a. A summary of the quality characteristics with their tolerances is depicted in Table 3.

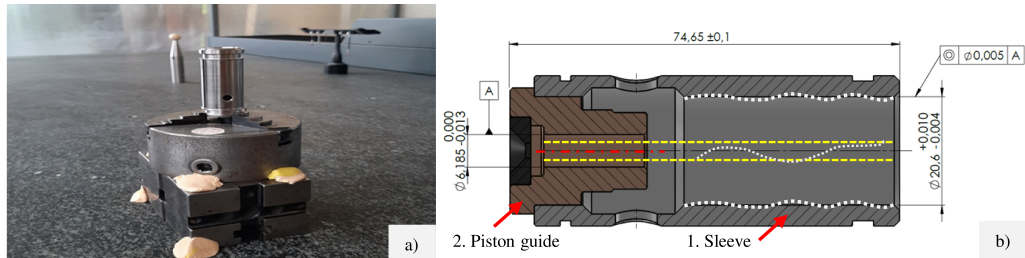


Fig. 1. a) Workpiece used in the study, positioned in the fixture on the table of the CMM and b) technical specification with the depicted coaxial tolerance between sleeve and piston guide.

Table 3. MPE values of the CMMs against the measured quality characteristics and respective accuracy ratios.

Quality characteristic	Nominal value	Tolerance	CMM 1MPEE	CMM 2MPEE
Diameter	20.6 mm	+0.01/−0.004 mm	$3.5\text{ }\mu\text{m} + 3 \times 10^{-6} \times L$	$1.6\text{ }\mu\text{m} + 3.5 \times 10^{-6} \times L$
Coaxiality	–	0.005 mm		
Length	74.65 mm	±0.1 mm		
Accuracy ratio for CMM 1	Interpretation for CMM 1	Accuracy ratio for CMM 2	Interpretation for CMM 2	
ARDiameter ≈ 3.9:1	Not acceptable	ARDiameter ≈ 8.4:1	Marginal	
ARCoaxiality ≈ 1.4:1	Not acceptable	ARCoaxiality ≈ 3.1:1	Not acceptable	
ARLength ≈ 53.7:1	Acceptable	ARLength ≈ 107.5:1	Acceptable	
CMM 1: Wenzel LH108 CMM 2: Wenzel XO87 NOTE: The accuracy ratio was calculated according to AS 13003:2015 [6].				

2.3.2. Coordinate measuring machines

The coordinate measuring machines utilized are a Wenzel LH108 and a Wenzel XO87. A Renishaw motorized head with a PH10M sensor was used on both machines, as well as a probe sphere with a diameter of 2 mm. The resolution of the machines was set to 0.1 μm . In Table 3, maximum permissible error values for length measurements and *maximum permissible error of indication* (MPEE) are given for both CMMs as well as previously mentioned nominal values with tolerances of the quality characteristics measured. *Accuracy ratio* (AR), in the Table 3 was calculated according to [6]. There is a clear difference in terms of accuracy between the selected coordinate measuring machines, and since the same sensor and probe tip were used on both machines, the assumption is that measurement errors and/or measurement uncertainty will be greater on the nominally “less accurate” machine.

For the diameter and length measurement, the ratio between tolerance and MPEE indicates that both CMMs operate within an acceptable or marginal range. Since the diameter is derived from coordinate measurements, the MPEE provides a reasonable approximation of measurement capability, although additional uncertainty may arise from the fitting procedure. For the coaxiality tolerance, the ratio between tolerance and MPEE is significantly below acceptable limits, since this is a critical feature, indicating that the measurement system is not capable of reliably evaluating this characteristic. Furthermore, it should be emphasized that MPEE is defined for length measurements and does not fully capture additional sources of uncertainty involved in geometric tolerances such as coaxiality, including form deviations, alignment procedures, and fitting algorithms.

2.3.3. Conducting a full factorial design of experiment

The collection and analysis of data were conducted in accordance with a full factorial *design of experiments* (DoE), utilizing the Minitab 21 and Origin Pro 2024 software packages. Measurements were performed for each quality characteristic on the 10 workpieces at two different temperature levels. Each measurement had 3 repetitions per condition, using a single measurement program on both coordinate measuring machines. That yielded 120 measurement results. The measurement program was created to be executed automatically, in the same way, for each workpiece. Before measuring a workpiece and after a change in temperature, calibration of all probe tip positions was performed using an appropriate spherical reference standard. Due to the advantages mentioned in [13], the experiment was carried out in a random order taking into account all variations. Since CMMs do not have identical accuracy characteristics throughout the entire measuring volume, during the experiment the fixture was placed in the same position for each workpiece [14]. The AS13003_1:2015 standard recommends performing a full factorial design of experiments where the variables are: (i) run-to-run variation on the same machine, (ii) variation between machines, and (iii) variation in ambient temperature. Table 4 shows the factors and levels used in the study.

Table 4. Full factorial DoE with factors and levels.

Factor	Levels	Values	
Workpiece (A)	10	1; 2; 3; 4; 5; 6; 7; 8; 9; 10	
Machine (B)	2	Wenzel LH108	Wenzel XO87
Temperature (C)	2	22°C	28°C

It often happens that a single measurement program is used on several different coordinate measuring machines in order to increase productivity when measuring a larger series of workpieces. However, it is necessary to determine the variability of the measuring systems used and their contribution to the total variation of the production process, especially in the case of strict requirements in the form of narrow tolerances of critical workpiece characteristics. In such cases, the variation caused by the measurement system may mask the size of the allowable deviation of the workpiece (the width of the tolerance zone) [15].

3. Results and discussion

3.1. Results for diameter, coaxiality and length from DoE

Before analysing the experimental results, additional analyses were carried out for outliers, homogeneity of variances, and normality testing of the data. The presence of outliers within the dataset can have a significant negative effect on further data analysis [16].

ANOVA for diameter, coaxiality, and length is shown in Tables 5, 6 and 7, respectively. For diameter, all observed factors as well as their interactions were statistically significant at the given 95% confidence level. All main factor effects proved to be statistically significant. While the workpiece factor dominated the total variation, the machine factor had the strongest influence on the response mean. Other contributors following machine factor were temperature factor.

Table 5. Analysis of variance for diameter.

Source	DF	Adj SS	Adj MS	Contribution	F-Value	P-Value
Model	39	0.025216	0.000647	99.66%	599.92	0.000
Linear	11	0.024869	0.002261	98.29%	2097.70	0.000
Workpiece	9	0.023817	0.002646	94.13%	2455.44	0.000
Machine	1	0.000902	0.000902	3.56%	836.94	0.000
Temperature	1	0.000150	0.000150	0.59%	138.84	0.000
2-Way Interactions	19	0.000215	0.000011	0.85%	10.51	0.000
Workpiece*Machine	9	0.000092	0.000010	0.36%	9.47	0.000
Workpiece*Temperature	9	0.000097	0.000011	0.38%	10.02	0.000
Machine*Temperature	1	0.000026	0.000026	0.10%	24.25	0.000
3-Way Interactions	9	0.000132	0.000015	0.52%	13.62	0.000
Workpiece*Machine*Temperature	9	0.000132	0.000015	0.52%	13.62	0.000
Error	80	0.000086	0.000001	0.34%		
Total	119	0.025302		100.00%		

Table 6. Analysis of variance for coaxiality.

Source	DF	Adj SS	Adj MS	Contribution	F-Value	P-Value
Model	39	0.004648	0.000119	39.09%	1.32	0.150
Linear	11	0.002206	0.000201	18.55%	2.22	0.021
Workpiece	9	0.001035	0.000115	8.70%	1.27	0.266
Machine	1	0.000279	0.000279	2.34%	3.08	0.083
Temperature	1	0.000893	0.000893	7.51%	9.86	0.002
2-Way Interactions	19	0.002033	0.000107	17.10%	1.18	0.294
Workpiece*Machine	9	0.000725	0.000081	6.09%	0.89	0.539
Workpiece*Temperature	9	0.001267	0.000141	10.66%	1.56	0.143
Machine*Temperature	1	0.000042	0.000042	0.35%	0.46	0.500
3-Way Interactions	9	0.000408	0.000045	3.43%	0.50	0.869
Workpiece*Machine*Temperature	9	0.000408	0.000045	3.43%	0.50	0.869
Error	80	0.007243	0.000091	60.91%		
Total	119	0.011890		100.00%		

Table 7. Analysis of variance for length.

Source	DF	Adj SS	Adj MS	Contribution	F-Value	P-Value
Model	39	0.587917	0.015075	99.97%	7086.79	0.000
Linear	11	0.584531	0.053139	99.40%	24981.22	0.000
Workpiece	9	0.583030	0.064781	99.14%	30454.18	0.000
Machine	1	0.001378	0.001378	0.23%	647.67	0.000
Temperature	1	0.000124	0.000124	0.02%	58.12	0.000
2-Way Interactions	19	0.002780	0.000146	0.47%	68.79	0.000
Workpiece*Machine	9	0.001779	0.000198	0.30%	92.91	0.000
Workpiece*Temperature	9	0.000930	0.000103	0.16%	48.59	0.000
Machine*Temperature	1	0.000071	0.000071	0.01%	33.45	0.000
3-Way Interactions	9	0.000606	0.000067	0.10%	31.63	0.000
Workpiece*Machine*Temperature	9	0.000606	0.000067	0.10%	31.63	0.000
Error	80	0.000170	0.000002	0.03%		
Total	119	0.588087		100.00%		

For the coaxiality, among the considered main factors and their interactions, only the temperature factor was found to be statistically significant, while all other p-values were higher than the significance threshold. The coaxiality tolerance in this case represents the most demanding quality characteristic with a tolerance zone width of $5\ \mu\text{m}$. Based on the calculated accuracy ratios from Table 4, neither coordinate measuring machine satisfies the commonly accepted 10:1 criterion for this characteristic. Despite this limitation, the analysis was intentionally carried out, as in practical industrial environments measurements are often performed using equipment that does not fully meet recommended capability requirements. Relying purely on the MPEE value will not give adequate results in such cases. Furthermore, it should be emphasized that relying solely on this value is insufficient for evaluating measurement capability in the case of geometric tolerances such as coaxiality. Consequently, the obtained results highlight the necessity of applying extended GR&R methodologies to achieve a more realistic assessment of measurement system performance.

For the verification of such a characteristic, controlled environmental conditions and a comprehensive evaluation of measurement uncertainty are generally required to ensure reliable conformity assessment with the specified tolerance.

In the analysis of the assembly's length, similar to the diameter, all observed factors and their interactions were statistically significant. The relative spread of the main factors is acquired from taking the square root of the mean squares, thus getting information about the standard deviation. The next step was plotting the distributions with the same mean values but different standard deviations. Mean squares are not pure variance components in this example. They include true variation, interaction effects, and residual errors.

Statistical significance of the considered factors suggests that the DoE was carried out soundly, and further analysis will shift its emphasis to the gage R&R. DoE served as a way for the proper collection of the data, and as an input for extended gage R&R.

3.2. Expanded GR&R Study

To evaluate the suitability of the selected CMMs for measuring the specified characteristics, an expanded repeatability and reproducibility study is conducted. Based on the previously described experimental design, the total variation of the study was divided into variation due to the workpiece

factor, the machine factor, and the temperature factor at two levels. Since this is a critical workpiece with strict dimensional and geometric tolerance requirements, in addition to the study variation data ($\%StudyVar$) and the contribution of individual factors to the total variation ($\%Contribution$), it is necessary to analyse the percentage of participation of individual factors and their interactions relative to the size of the tolerance zone ($\%Tolerance$). The $\%Tolerance$ column is also referred to as the precision-to-tolerance ratio – PTR. Figure 2 graphically shows the p -values from the ANOVA tables obtained from the GR&R analysis and compares them with the previously obtained p -values from the analysis of the experimental design. The significance threshold was $\alpha = 0.05$.

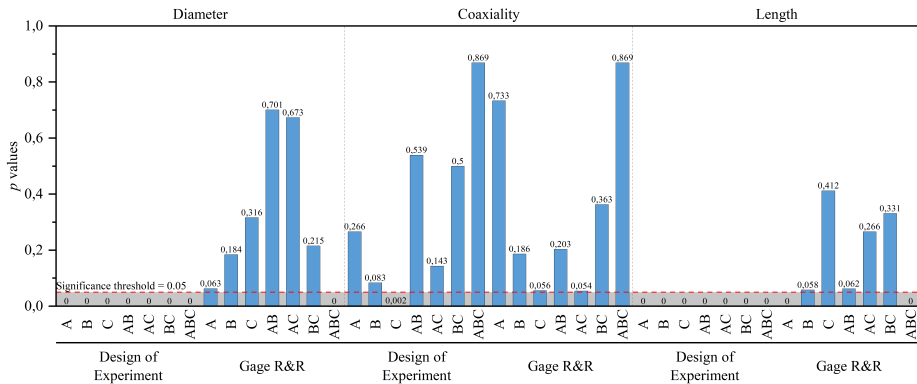


Fig. 2. Graphical review of the p -values from ANOVA tables obtained through general full factorial design of experiment and expanded gage R&R study

When comparing the p -values from Fig. 2, a certain discrepancy in the obtained values can be observed. Most factors and their interactions are statistically significant when considering the ANOVA from the classical experimental design, while in the ANOVA obtained from the GR&R study most factors and interactions are not statistically significant. With respect to the quality characteristics, the most similar p -values were obtained for the characteristic of length, while for the coaxiality, the greatest discrepancy was observed.

Explanation of the observed discrepancy between p -values acquired through Design of experiment and extended GR&R fundamentally lies in interpretation. In the DoE analysis, factors are treated as fixed and the objective is to determine their significance based on their influence on the response *i.e.*, determination of the factor main effects and their interactions. This is done by calculations of mean responses across factor levels and subsequent comparisons. In contrast, GR&R is based on the random effect model. Emphasis in that case is on the variances of the individual components and their contribution to the overall study variation or ratios against the specific tolerance. As stated by Montgomery [17], the procedure for the computation of the random effects models is identical to the fixed-effects case. However, these approaches are quite different, because GR&R (random-effect) assumes that factors are random samples from a population and on the other hand the fixed-effect model conclusions can only be applied to the specific factor levels used in experiment.

In the present study, this implies that while the DoE analysis may detect differences in mean responses between workpieces or measurement conditions, the GR&R results indicate that the part-to-part variation is small relative to the measurement system variation. Consequently, the measurement system lacks sufficient resolution to reliably distinguish between individual workpieces.

The histograms in Fig. 3a display the percentage of the total measurement system variation relative to the overall study variation, *i.e.*, the size of the tolerance zone for the observed quality characteristics, and Fig. 3b, the contribution of the variation components to the overall study. Considering the minimum requirements set by the aviation standard AS13003_1:2015, the measurement system variation is acceptable only for the measurement of the workpiece length, since the components of the total measurement system variation, *i.e.*, repeatability and reproducibility, are below the threshold value of 30%. However, the 30% limit applies in cases where the quality characteristic is not of critical importance. From Table 1, the values for the total measurement system variation should be less than or equal to 10%. The most unfavourable results were obtained in the measurement of coaxiality, which was expected given the size of the tolerance zone.

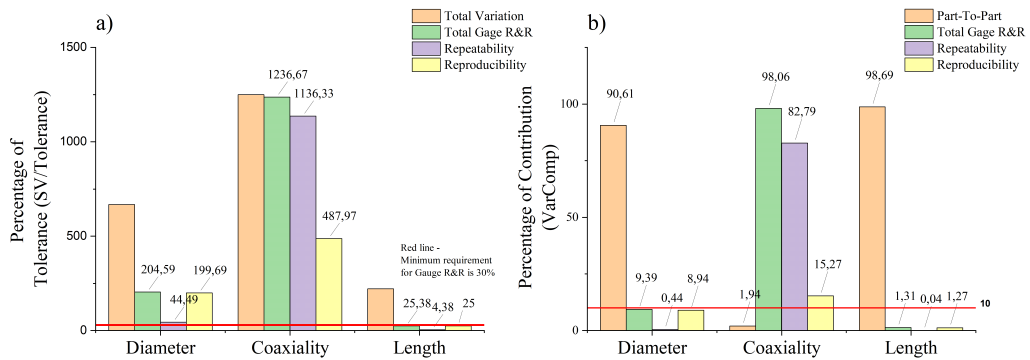


Fig. 3. Summarized results from GR&R histograms regarding a) %Tolerance and b) %Contribution for the quality characteristics.

Looking at the histogram of the contributions of individual components to the total variation of the study in Fig. 3b, the selected measurement systems achieved satisfactory results for the diameter and length characteristics. Most of the variation in the study is attributed to part-to-part variation. In contrast, the inadequacy of the measurement system for verifying coaxiality is also reflected here. The total measurement system variation amounts to slightly more than 98 %, with the larger portion accounted for by measurement system repeatability. Since two measurement systems were included in the study, repeatability variation represents the variation within the machines, while reproducibility corresponds to the variation between the selected measurement systems. The relatively low value of reproducibility (still unacceptable) is explained by the fact that the same fixture and the same measurement strategy were used on both coordinate measuring machines.

The comparison of contribution percentages obtained from the DoE and the extended GR&R approach provides additional insight into the influence of the measurement system on the observed variation. Since the primary objective is to assess measurement system adequacy, the contribution of the CMMs obtained from ANOVA (Tables 5–7), *i.e.* the machine, the main factor only, is compared with the Total GR&R values derived from variance components. The results are presented in Fig. 4. For diameter and length, both approaches indicate consistent trends where the measurement system contribution remains relatively low compared to the part-to-part variation, although the numerical values differ by less than 10 %. In contrast, for coaxiality, the DoE results suggest a relatively low contribution of the machine factor, while the GR&R analysis indicates that the dominant source of variation originates from the measurement system. This inconsistency can be attributed to the large residual error in the ANOVA model (approximately 61%, see Table 6), which masks the true contribution of the measurement system.

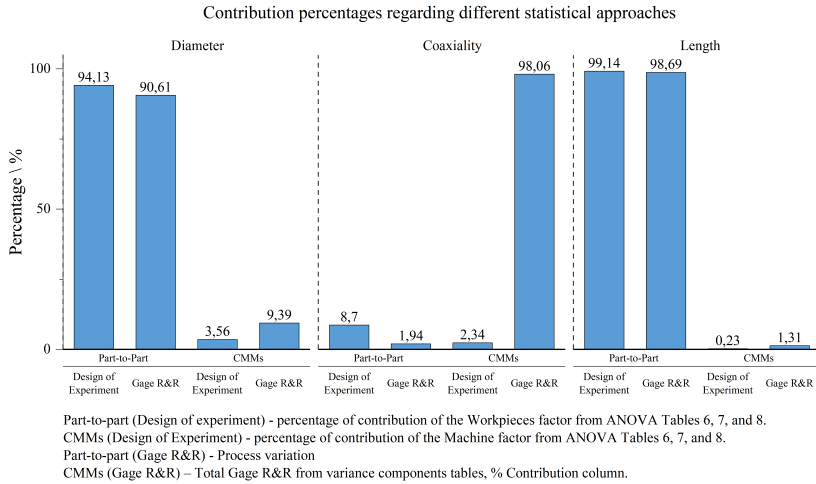


Fig. 4. Comparison of contribution percentages obtained from the DoE and the extended GR&R approach.

Figure 5 shows box-plot diagrams for the observed quality characteristics. The median line connects adjacent box-plots. For the diameter characteristic, greater dispersion of results was observed for the Wenzel XO87, on workpieces 5, 6, and 7. For coaxiality, significant differences in

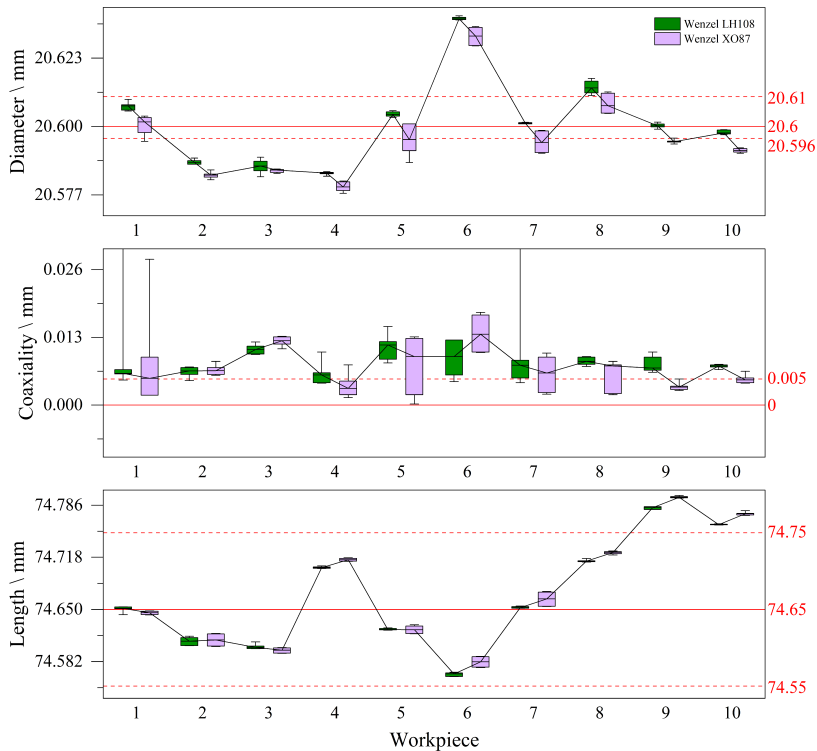


Fig. 5. Box-plot diagrams for the quality characteristics regarding CMM performance based on different workpieces.

result dispersion are noted for workpieces no. 1 and 7 with the Wenzel LH108, while for the length measurement, qualitatively, no major differences exist in the results produced by the measurement systems across workpieces.

Figure 6 shows the dispersion of results and medians for individual quality characteristics with respect to environmental conditions expressed through the temperature factor. These diagrams include nominal values and tolerance limits for each observed characteristic. For the diameter of the workpiece, the Wenzel LH108 in most cases provides results that meet the technical specification, while this is not the case for the Wenzel XO87. For the coaxiality characteristic, only at the higher temperature level does the Wenzel XO87 provide results within the specification limits. Interestingly, at this temperature level, both measurement systems show less result dispersion compared to the temperature closer to the reference value of 20 °C.

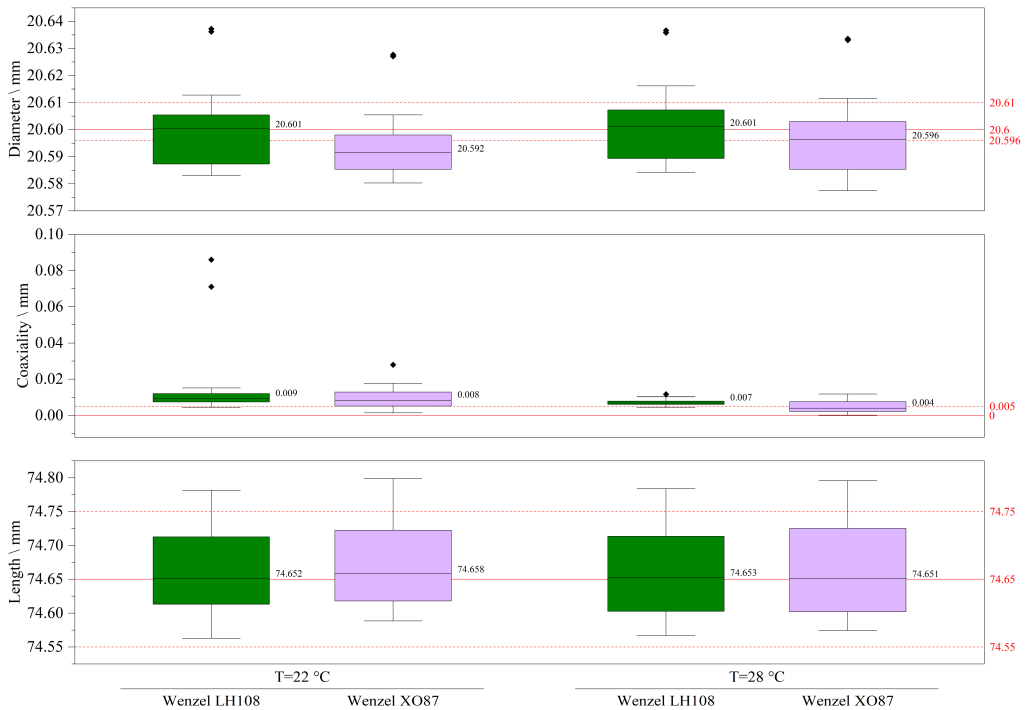


Fig. 6. Box-plot diagrams for the quality characteristics regarding CMM performance based on different environmental conditions.

The ANOVA study conducted within the GR&R provides insight into interactions, shown in Fig. 7. The interaction diagram of measurement systems depending on the workpiece, for diameter, Fig. 7a, shows that interactions are non-existent, since the broken lines do not intersect from one workpiece to another. Generally speaking, the Wenzel XO87 yields lower mean values compared to the Wenzel LH108, for each workpiece. The interaction diagram for the coaxiality, Fig. 7b, shows the crossing of mean value lines up to workpiece no. 7, confirming the existence of interactions among the observed factors. In the interaction diagram for length, Fig. 7c, there is no interaction between factors.

The general conclusion from the conducted experiment is that industrial coordinate measuring machines whose maximum permissible error is greater than 1 μm are not adequate for the evaluation of quality characteristics whose tolerance values are in the micrometre range.

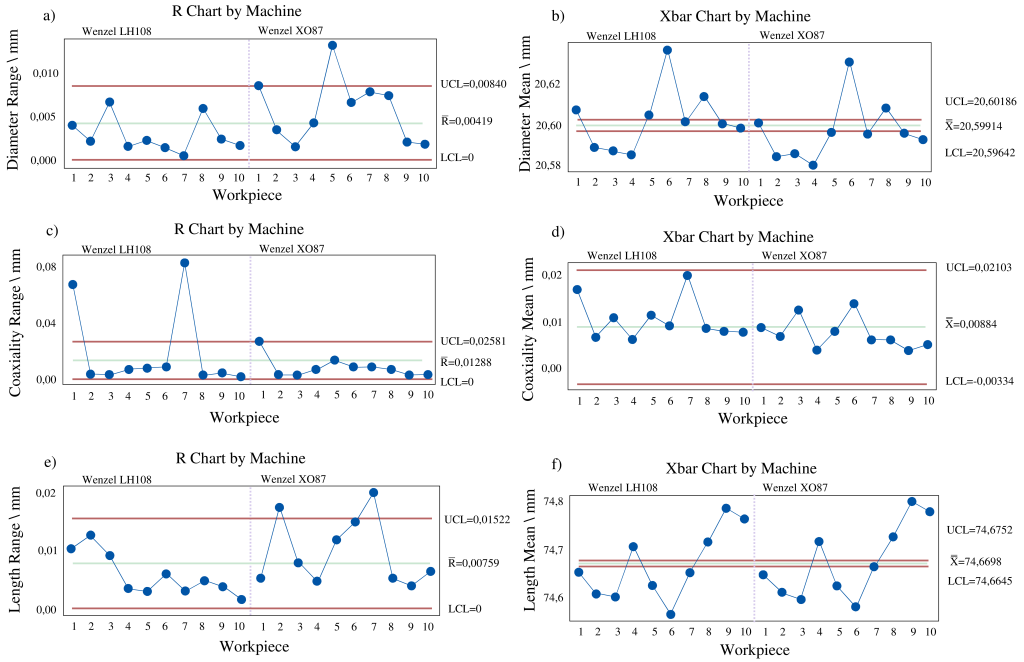


Fig. 7. Interaction plots depicting the machine-workpiece relationship for different quality characteristics.

Lastly, several distinct categories should be mentioned. NDC represents the ratio of part-to-part variation to the total measurement system variation in the study. It indicates how well the measurement system is able to “distinguish” workpieces, *i.e.*, to classify them into different categories. Results for NDC support earlier conclusions about the inadequacy of selected coordinate measuring machines to evaluate two out of three characteristics discussed in this study. For the diameter characteristic, NDC is 4, for the length it is 12, while for coaxiality it is 1. Requirements for the minimum NDC number are 5, according to [5].

4. Conclusions

Assessing the capabilities of a measurement system to perform certain measurement task adequately is a key procedure for ensuring the compliance of manufactured products with their specifications. In this study, an extended GR&R methodology, supported by a structured DoE, was applied to evaluate the performance of two CMMs in measuring diameter, length, and coaxiality.

The results demonstrate that measurement system capability strongly depends on the type of quality characteristic and its associated tolerance. Both CMMs satisfied the acceptance criteria only for the measurement of length, while their performance was inadequate for diameter and coaxiality when evaluated against tolerance-based criteria. In particular, coaxiality, as the most demanding characteristic, exhibited dominant measurement system variation, primarily due to repeatability issues and sensitivity to environmental conditions.

A comparison between DoE-based ANOVA and extended GR&R revealed important differences in interpretation. While the fixed-effect DoE approach identified statistically significant factors based on mean differences, the GR&R analysis, based on random effects, provided insight into

variance components and measurement system contribution. This distinction proved critical, especially for coaxiality, where the DoE model underestimated the influence of the measurement system due to large residual error.

The findings highlight that reliance on a single statistical approach may lead to incomplete or misleading conclusions regarding measurement system adequacy. For complex geometrical characteristics and tight tolerances, extended GR&R analysis provides a more realistic assessment, particularly when measurement uncertainty and environmental influences are significant.

Based on the obtained results, it can be concluded that for reliable measurement of characteristics with tolerances in the micrometre range (e.g., coaxiality of 5 μm), coordinate measuring machines should exhibit maximum permissible errors significantly below 1 μm , ensuring compliance with the 10:1 accuracy ratio criterion [6]. This implies the need for accurate and precise CMMs, stationed in laboratory-grade environmental conditions.

References

- [1] Wojtyła, M., & Płowucha, W. (2025b). Measurement models – theory and practice of uncertainty evaluation of geometrical deviation measurements. *Metrology and Measurement Systems*, Vol. 32 (2025) No. 3, 1–21. <https://doi.org/10.24425/mms.2025.154671>
- [2] Stępień, K., Świdorski, J., & Makiela, W. (2025). The importance of measurement uncertainty during gauge block calibration in determining compliance with specifications. *Metrology and Measurement Systems*, Vol. 32 (2025) No. 2, 1–15. <https://doi.org/10.24425/mms.2025.154338>
- [3] Wozniak, A., Krajewski, G., & Byszewski, M. (2018). A new method for examining the dynamic performance of coordinate measuring machines. *Measurement*, 134, 814–819. <https://doi.org/10.1016/j.measurement.2018.12.041>
- [4] Štrbac, B., Rodić, D., DeliĆ, M., Savković, B., & Hadžistević, M. (2021). Investigation of Functional Dependency between the Characteristics of the Machining Process and Flatness Error Measured on a CMM. *Measurement Science Review*, 21(6), 158–167. <https://doi.org/10.2478/msr-2021-0022>
- [5] Chrysler Group LLC, Ford Motor Company, & General Motors Corporation. (2010). *Measurement Systems Analysis: Reference Manual* (4th ed.). Automotive Industry Action Group (AIAG).
- [6] SAE International. (2015). AS13003: *Measurement Systems Analysis Requirements for the Aero Engine Supply Chain*, <https://doi.org/10.4271/AS13003>
- [7] Minix, J., Chapman, H., Joshi, N., & Zargari, A. (2016). An Investigation of Measurement Uncertainty of Coordinate Measuring Machines (CMMs) by Comparative Analysis *The Journal of Technology Studies*, 42(1), 54–63. <https://doi.org/10.21061/jots.v42i1.a.5>
- [8] Dev, A., Jha, P. K., & Parkash, V. (2018). Effect of automation on GRR in tactile measurement. *MAPAN*, 33(4), 449–458. <https://doi.org/10.1007/s12647-018-0269-1>
- [9] Shirodkar, A., & Rane, S. (2021). Evaluation of Coordinate measuring machine using Gage Repeatability & Reproducibility. *International Journal of Systems Assurance Engineering and Management*, 12(1), 84–90. <https://doi.org/10.1007/s13198-020-01050-2>
- [10] García, A. C., & Del Río, A. G. (2013). Number of distinct data categories and gage repeatability and reproducibility. A double (but single) requirement. *Measurement*, 46(8), 2514–2518. <https://doi.org/10.1016/j.measurement.2013.04.065>
- [11] Razumić, A., Runje, B., Alar, V., Štrbac, B., & Trzun, Z. (2025). A review of methods for assessing the quality of measurement systems and results. *Applied Sciences*, 15(17), 9393. <https://doi.org/10.3390/app15179393>

- [12] Montgomery, D. C. (2020). *Introduction to Statistical Quality Control* (8th ed.). John Wiley & Sons.
- [13] Burdick, R. K., Borror, C. M., & Montgomery, D. C. (2003). A review of methods for measurement systems capability analysis. *Journal of Quality Technology*, 35(4), 342–354. <https://doi.org/10.1080/00224065.2003.11980232>
- [14] Štrbac, B., Ačko, B., Havrlišan, S., Matin, I., Savković, B., & Hadžistević, M. (2020). Investigation of the effect of temperature and other significant factors on systematic error and measurement uncertainty in CMM measurements by applying design of experiments. *Measurement*, 158, 107692. <https://doi.org/10.1016/j.measurement.2020.107692>
- [15] Saikaew, C. (2018). An implementation of measurement system analysis for assessment of machine and part variations in turning operation. *Measurement*, 118, 246–252. <https://doi.org/10.1016/j.measurement.2018.01.008>
- [16] Abduljaleel, M., Midi, H., & Karimi, M. (2019). Outlier detection in the analysis of nested gage R&R, random effect model. *Statistics in Transition New Series*, 20(3), 31–56. <https://doi.org/10.21307/stattrans-2019-023>
- [17] Montgomery, D. C., & Runger, G. C. (2010). *Applied Statistics and Probability for Engineers* (7th ed.). Hoboken, NJ: Wiley.



Miloš Ranisavljev received his B.Sc. degree in 2020 and his M.Sc. degree in 2021 from the Faculty of Technical Sciences, University of Novi Sad, Serbia. He is currently employed as a teaching assistant in the field of metrology and quality. His research interests focus on application of coordinate measuring systems. In addition to his teaching responsibilities, he is pursuing his Ph.D. degree with research focused on dimensional metrology using computed tomography.



Branko Štrbac received his B.Sc. and M.Sc. degrees from the Faculty of Technical Sciences, University of Novi Sad, in 2009, and his Ph.D. degree from the same faculty in 2017. He is currently a professor at the Faculty of Technical Sciences, University of Novi Sad, where he teaches courses related to production metrology and serves as a Ph.D. supervisor. His main research interests include precision measurement technology, coordinate metrology, and precision measuring instruments.



Predrag Janković graduated in mechanical engineering, received his PhD degree from the Department of Production Engineering in 2009. He is a full professor at the Faculty of Mechanical Engineering, University of Niš, Serbia, and, also, head of the Laboratory for Engineering Metrology. His teaching courses and research areas include metrology, quality management, measurement systems analysis, and optimization of production processes.



Sara Havrlišan received the Ph.D. degree in 2017. She is currently an assistant professor at the Faculty of Mechanical Engineering, University of Slavonski Brod, Croatia. Her research interests include optimization, thermal spraying, design of experiments, Ni-based self-fluxing alloys, flame spraying, and regression modelling.



Branislav Dudić is a faculty member at the Faculty of Management, Comenius University Bratislava. His research interests include industrial engineering, business digital transformation, human resource management, marketing management with a focus on implementation of marketing strategies, and investments in both the private and public sectors within the European Union.



Miodrag Hadžistević received the B.Sc. degree from the Faculty of Technical Sciences, University of Novi Sad, in 1992, the M.Sc. degree in 1999, and the Ph.D. degree in 2004 from the same faculty. He is currently a full professor at the Faculty of Technical Sciences, University of Novi Sad. His research interests include manufacturing metrology, design of experiments (DoE), and occupational safety and health.