APEM journal
Advances in Production Engineering & Management
Volume 19 | Number 2 | June 2024 | pp 209–222
https://doi.org/10.14743/apem2024.2.502
ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Unsupervised machine learning application in the selection of measurement strategy on Coordinate Measuring Machine

Štrbac, B.[a,*], Ranisavljev, M.[a], Orošnjak, M.[b], Havrlišan, S.[c], Dudić, B.[d,e], Savković, B.[a]

[a]University of Novi Sad, Faculty of Technical Sciences, Department of Production Engineering, Novi Sad, Serbia
[b]University of Novi Sad, Faculty of Technical Sciences, Department of Industrial Engineering and Management, Novi Sad, Serbia
[c]University of Slavonski Brod, Mechanical Engineering Faculty, Slavonski Brod, Croatia
[d]Comenius University Bratislava, Faculty of Management, Bratislava, Slovakia
[e]Faculty of Economics and Engineering Management, University Business Academy, Novi Sad, Serbia

A B S T R A C T

It is indisputable that some type of coordinate measurement system (CMS) is generally used to assess the quality of dimensional and geometric characteristics. Considering the required accuracy, flexibility, and speed of measurement, a CMM with a scanning sensor may offer the best performance. These measurement systems are very complex, and many factors affect the reliability of the measurement results. A Metrologist's choice represents the greatest variability in the measurement strategy. Previous research has shown that the measurement results can be changed up to 100 % by choosing a different measurement strategy when evaluating the form error. This paper conducts a detailed study of the impact of the measurement strategy on the cylindricity error when measuring eleven workpieces with the same nominal characteristics, but different real characteristics described by roughness and the reference value of cylindricity. To examine the influence and importance of certain factors and their levels, design of experiment (DoE) and unsupervised machine learning techniques of PCA (Principal Component Analysis) and Multiple Correspondence Analysis (MCA), were used. The results suggest that depending on the real geometry of the workpiece, different factors with different percentages influence the output characteristic. The objective was to choose a uniform measurement strategy when measuring cylindricity on the CMM, while the prior information about the actual geometry of the workpiece is lacking.

A R T I C L E   I N F O

Keywords:
Coordinate Measuring Machine (CMM);
Measurement strategy;
Accuracy;
Principal component analysis;
Multiple correspondence analysis;
Unsupervised learning

**Corresponding author:*
strbacb@uns.ac.rs
(Štrbac, B.)

Article history:
Received 17 April 2024
Revised 27 June 2024
Accepted 30 June 2024

1. Introduction

Quality assessment of dimensional and geometric product specifications in the field of precision engineering is generally performed on some type of coordinate measuring system (CMS). In the last two decades, various versions of CMS have appeared, such as coordinate measuring machine with contact and non-contact sensors, photogrammetry, measuring arms, non-contact (optical and laser) scanners, industrial CT, etc. [1]. Each type has its advantages and disadvantages, which are primarily reflected in the accuracy of measurement, flexibility, the amount of information collected from the workpiece (points), inspection time, price, etc. [2, 3]. Considering all the above requirements, perhaps a coordinate measuring machine (CMM) equipped with a contact-scanning sensor is the most optimal solution for the verification of the narrow specification requirements imposed by the functional requirements of the workpieces. However, CMMs are very complex measurement systems and there are numerous influencing factors and their interactions that contribute to a large range of dispersion of results, which can be correlated with measurement uncertainty [4]. It seems that the assessment of measurement uncertainty remains within the framework of laboratory measurements for calibration and even theory for practical industrial measurements. So far developed methods for measurement uncertainty assessment such as GUM or series of ISO 15530 standards either have many limitations and/or require calibrated parts, repeated measurements, development of simulation methods, etc. [5-11]. For this reason, measurement uncertainty is generally not given with the results of practical measurements, although it is the only indicator of the quality of the measurement results, and the maximum permissible error (MPE) is often taken as an indicator of the quality of the measurement. Also, repeatability and reproducibility studies are not effective with CMMs because they are automated measuring systems.

In the total uncertainty budget of the CMM, e.g. according to ISO 15530-3, which is considered a reference standard, it figures the uncertainty of calibration of the workpiece, repeated measurements, temperature effects, while the measurement strategy is not considered because it is the same for all 20 repeated measurements required by the standard [6]. However, a certain issue is the selection of the measurement strategy because the specification operators are not complete, and the selection of the measurement strategy is left to the operator [12]. The results of some measurement strategies may show compliance with the specification while others may not. Also, the actual dimensions of the parts from the production that are measured always remain unknown, so no judgment can be made about the "good" or "bad" chosen measurement strategy. Under the measurement strategy of CMM can be considered the number and position of measurement points on the observed feature, the criteria for creating a fitted geometry, the choice of the tip of the measuring probe and, if they are scanning contact sensors, the scanning speed and filtering. Yes, current research has shown that these parameters in interaction with the form deviation of the measured primitive really affect the measurement uncertainty [13, 14]. Uncertainty factors originating from geometrical errors of the CMM, and measurement sensor cannot be controlled because they are inherent to the measuring instrument and will always be present in the overall measurement result, but the uncertainties originating from the choice of measurement strategy can be greatly influenced by choosing the optimal measurement strategy for ordered measurement task. In papers [15, 16], the selection of the optimal number of points was made for obtaining measurement values that are close to "actual". Certainly, the number of points tending to infinity would be the best for describing actual geometry, but it is not economically justified, and in these publications, it was shown that depending on the processing procedure, the number of optimal points also depends on the example of flatness measurement. These studies have shown that the choice of strategy when evaluating the form error depends primarily on the processing procedure used to obtain the workpiece, and the value of the mean arithmetic roughness was used as input information.

The aim of this work is not to evaluate the measurement uncertainty but to demonstrate how much the choice of measurement strategy with contact scanning sensors can affect the scattering of results when measuring cylindricity on workpieces that are of the same nominal dimensions but were manufactured with different processes, different settings and are characterized by deviations in the quality of the surface. To examine the variability of measurement results due to a change in measurement strategy, the design of the experiment was used. In addition, using dimensionality reduction tools, i.e. Principal Component Analysis (PCA) and Multiple Correspondence Analysis (MCA), the optimal solution of measurement strategies has been determined by considering all objects with same nominal dimensions but with deviations in geometry.

2. Methodology

2.1 Experimental setup

To choose the optimal strategy for measuring cylindricity on a CMM in scanning mode, eleven workpieces were manufactured. All workpieces had the same nominal dimensions, an outer diameter of 12 mm and a height of 14.9 mm and are made by different technological procedures, grinding and turning, with different cutting parameters to achieve variability in terms of form error and roughness of the processed surface, Fig. 1. The samples selected in this way cover a wide range of imperfections of real workpieces that can be found in practice when it comes to cylindrical workpieces. The measure of the diversity of the workpieces is the difference in the mean arithmetic roughness Ra and the reference (actual) cylindricity $CYLt_{ref}$, which was obtained by measuring the workpieces on the Roundtest, Table 1. The results obtained on the Roundtest were used as target (desired) values of cylindricity and were used to optimize the measurements on CMM. Roundtest is a measuring instrument specialized for measuring form error, primarily circularity, and cylindricity, and is used to verify very narrow specification limits and generally has a much higher accuracy than coordinate measuring machines. For the purposes of this experiment, Roundtest RA-2200AS Mitutoyo, radial rotational accuracy $(0.04+6H/10000)$ µm was used; H is measuring height in mm and axial rotational accuracy $(0.04+6X/10000)$ µm; X is the distance from the rotation center in mm. The workpieces were measured with a measuring probe of 0.75 mm in radius, with twelve concentric circles (each circle contained 7000 points), and a Gaussian-50%Low [50 UPR] filter was used. Fig. 2 shows the measurement report from the roundtest for cylinder 7.

Following the roundtest measurements, workpieces were measured on a coordinate measuring machine Carl Zeiss Contura G2 (Oberkochen, Germany), with the RDS rotating head and VAST XXT tactile sensor. The helix measurement strategy was used during the experiment, applying the scanning method. The maximum permissible error of the machine in the scanning mode is $MPE_{THP/\tau} = (2.2 + L/330)$ µm, defined according to DIN EN ISO 10360-4:2001. The measurements were conducted in the temperature-controlled environment of 20 ±0.5 °C. At about 80 % of the total height of the workpieces, points were sampled for the creation of substitute geometry, using the minimal zone method (MZ) and estimation of cylindricity.

Based on previous research and experience data, it was observed that the factors like diameter of the tip of the measuring probe, scanning speed, number of sampled points and filtering have the most influence on the error of cylindricity and are related to the measurement strategy on the CMM. To quantify the influence of individual factors on the variation of the output characteristics, the Methodology of the response surface - Box Behnken experiment design was used. The factors and levels used are shown in Table 2. The choice of factor levels was based on experience from practical measurements and the capabilities offered by CMM hardware and software.

Table 1 Measured values of the roughness parameter Ra and cylindricity, both in µm

Cylinder number	1	2	3	4	5	6	7	8	9	10	11
R_a	0.28	1.27	1.57	1.80	3.69	5.21	4.49	7.01	11.62	11.56	15.37
$CYLt_{ref}$	5.57	9.40	16.86	11.11	25.07	24.12	21.90	29.22	46.40	46.01	116.96

Fig. 1 Sample of workpieces

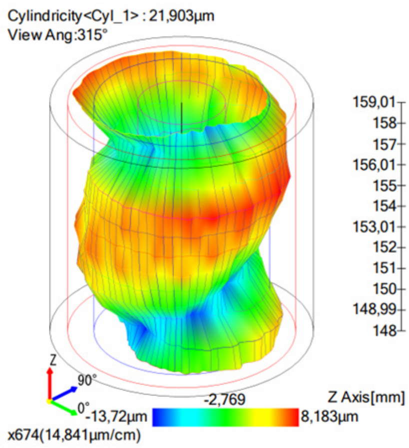


Fig. 2 Cylindricity result on Roundtest

Table 2 Factors used in Box-Behnken DoE in CMM measurement

Factors	Levels		
	-1	0	1
Measuring speed	5	35	100
Number of points	25	75	125
Stylus type diameter	2	3	5
Filter	Without	Cut-off 0.8	Cut-off 8

Measuring speed significantly impacts the measurement accuracy of CMMs. High scanning speeds can introduce dynamic flexibilities that affect measurement accuracy [18]. Ali *et al*. explored the impact of scanning speed and fitting algorithms on measurement results for cylindrical standards. They investigated five different scanning speeds in combination with many available fitting algorithms and concluded that for roundness measurements of the standard, the maximum inscribed fitting technique and 2 mm/s scanning speed give the lowest potential for errors [20]. Urban *et al*. investigated the influence of scanning speeds in the range from 2 up to 40 mm/s and the average deviation from reference size and roundness of gage rings. Evaluation of the experiment revealed that the scanning speed value used in combination with different radii of measured circular path quite fundamentally affects the accuracy and precision of measurement because increasing speed leads to increasing deviations due to the inertia of CMM construction. This increase is more pronounced for smaller diameters. For the roundness measurement, the same trend can be detected. Up to 6 mm/s speed, deviations are approximately the same for different ring gages, but above this speed, significant differences are observed [21]. Stępień [19] also investigated the influence of selected measurement parameters and strategy on the evaluation of form deviation values. Flatness and straightness were measured with five different scanning speeds and a number of sampling points. For the straightness evaluation, he chose a cylindrical workpiece. Like in the previous research, results indicate that form deviation increases along with the increase of scanning speed, without changing the number of sampling points.

The CMM inspection is based on a discrete measurement strategy, which is inherently an approximation process. If the sample size is infinite, the error in approximation approaches zero, and whenever the sample size is finite, the error must be non-zero. Since the number of points must be finite and much smaller than the number of points required by the ideal specification operator, the user of a CMM must plan the right measurement strategy to achieve the highest possible measurement accuracy [22]. The set of sample points is assumed to represent the feature being measured. These sample points are used to create the substitute profile for the feature being measured. The substitute profile is compared with the design profile to determine conformance to the specification. Intuitively, an increased sample size could lead to better characterization of the profile, however, the sample size is often limited by cost and time constraints. Thus, for a given sample size, the measurement strategy used must determine the locations and number of these measurement points such that the actual form may be effectively characterized [23].

Dimensional measurements performed by a CMM typically include all measured data points in an "as collected" manner. In other words, the raw, measured data points were directly analyzed without any pre-processing. However, as data densities have increased there has been an increased awareness that there is a great deal of "noise" in these "high density" data sets. This "noise" can be the result of such things as the surface roughness of the component being measured or due to errors in the measurement system like vibration. In many cases, it is desirable to filter out this "noise" to arrive at a more stable data set that is perhaps more indicative of the attributes that are to be assessed. In this regard, filtering methodologies have been put forth. These methodologies have been very common in the context of roughness and roundness measurement and are now making their way into other dimensional metrology applications [24]. According to Weckenmann [13], one factor in measurement strategy that has been largely ignored and has a significant impact on the results of coordinate measurements is the filtration of measurement points. From the experiments conducted on the geometrically ideal cylinder and the cylinder subject to deviation (mean value and standard deviation) effect of different cut-off values of the

Gaussian filter on the parallelism deviation is evident. For the range from 2000 up to 5000 sampled points, the lowest deviation, on both cylinders is acquired with the use of λc = 8.0 mm cut-off value, and the largest deviation is present where there is no filtration of sampled points.

The Stylus tip measurement diameter introduces mechanical filtration for sampling deviations from the workpiece. This factor is particularly important when measuring roundness because it filters out the effect of roughness and waviness on the form error [9]. Due to the finite size of the stylus, it is unable to access some deep valleys, and thus the measured surface is not the true surface but an approximate one, i.e. the mechanical surface. Normally the rougher the surface is, the more obvious the mechanical filtering effect will be. For the measurement of workpiece surfaces, the distortions caused by the tip mechanical filtration effect noticeably influence measurement accuracy [25].

2.2 Principal component analysis

Principal Component Analysis (PCA) is a multivariate statistical and unsupervised machine learning technique used for dimensionality reduction. Namely, the PCA transforms original set of variables p into a new set of uncorrelated variables by ordination, commonly referred to as principal components (PCs). These PCs are ordered by the amount of variance they capture from the original dataset. Thus, the PC1 accounts for the highest variation present, followed by PC2,…,PC*n*. It is important to note that ordination of PCs is under the constraint that they orthogonal on each previous PC. Performing the PCA analysis is conducted through a linear algebra and eigenvalue decomposition of a covariance and a correlation matrix. Let us consider that the original dataset X of dimensions $n \times p$, where n represents the number of observations, in our case measurements, while p represents the number of variables, in our case cylinders. Next, the PCA then converts the original set of variables into a set of uncorrelated variables (PCs), which capture the maximum amount of variance [26].

Explaining the procedure analytically, the first step of PCA is to perform standardization of the original dataset by z standardization. The underlying reason is that some variables may have different scales and range, which in turn, influence the PCA results. The common practice is to standardize dataset having the mean of 0 and a standard deviation of 1. The standardized dataset or matrix Z is then expressed as:

$$Z = \frac{X - \mu}{\sigma}, \tag{1}$$

where μ and σ are mean and standard deviation of the original dataset, respectively. After standardization, the covariance matrix C is computed to gain insight into variance and covariance of shared variables. The covariance matrix is then given by

$$C = \frac{1}{n-1} Z^T Z, \tag{2}$$

where Z^T is just transposed Z data matrix. After obtaining covariance matrix C, the eigenvalue decomposition is performed. In this step, the PCA computes eigenvalues and eigenvectors of the covariance matrix. The eigenvectors represent the direction of the maximum variance (PCs), while eigenvalues depict magnitude of the respected PC variance. The eigenvalue is expressed as:

$$Cv_i = \lambda_i v_i, \tag{3}$$

where v_i is the eigenvector and the λ_i is the eigenvalue of the covariance matrix. The eigenvectors obtained are ordered by decreasing eigenvalues. The first k principal components account for most of the variance in the data, where k is chosen based on the cumulative percentage of variance explained, in our case using a scree plot. The ordination into a new space is projected by k-dimensional feature space (where $k < p$) based on the selected eigenvectors that form new feature space via PCs. This transformation is obtained by

$$Y = ZV_k, \tag{4}$$

where V_k is the projected data matrix containing k selected (by order) eigenvectors, while Y represents dataset in the reduced feature space. It is important to note that simplification, i.e. transformation into lower-dimensional space is performed through linear combinations of uncorrelated principal components. Such transformation, although provides a more insightful understanding of the underlying structure of the dataset, is linearly dependent. Lastly, after obtaining lower dimensional space, our idea is to allocate (potential) clusters that share the most variance with projected reference measurement $x_{i, ref}$ from which we can obtain the factor set. However, instead of using additional machine learning algorithms, such as k-means, Agglomerative Hierarchical Clustering, or similar distance-based metrics, we turn attention to allocating the similarities between groups by using correspondence analysis techniques. The underlying reason for such an approach is that the method deals with categorical data, unlike clustering techniques that require numerical and continuous data for performing distance-based analysis (e.g., Euclidian distance).

2.3 Multiple correspondence analysis

The Correspondence Analysis (CA) is another dimensionality reduction technique that is, unlike PCA, used for categorical variables. Namely, the CA is performed by transforming original dataset of categorical variables into contingency tables. Thus, it accounts for count data between, for instance, two categorical variables. After obtaining frequencies via contingency table, the CA then projects the association between specific categories using Pearson's χ^2 distance metric [17]. The contingency table(s) consider I categories (rows) of specific variable p_i, where i = 1, 2,…,I, and J categories (columns) of variable p_j, where j = 1, 2,… ,J. Thus, the value in the dataset $x_{i,j}$ corresponds to the value of a variable with i rows and j columns, which consits of l instances. This can usually be performed by pivot tables accounting for frequencies between two categorical variables. Unlike proposed simple CA, which is constrained to the use of only two variables, the Multiple Correspondence Analysis (MCA) is able to perform singular value decomposition by using multiple categorical variables.

The process of MCA includes decomposition of matrix of indicator variables formed by transforming original set of categorical variables into a binary set of variables. In this case, however, the MCA computes only column profiles, i.e. J categories. Thus, the MCA performes weighted-PCA of the indicator matrix L, where $L = I \times H$. The levels of categorical variables p_h, such that h = {1, 2,…,H}, which is a binary representation of column profiles, i.e. categorical levels of variables, while I is number of observations i = 1, 2,…,I. If the number of categories are C_h, where h = 1, 2,…,H then there is C_h – 1 dimensions. Performing the calculation of the correspondence matrix P, from an indicator matrix L, is derived by dividing each element from the number of observations I to obtain relative frequency, i.e. probabilities as

$$P = \frac{1}{I} L^T L \,, \tag{5}$$

where L^T is the transposed indicator matrix L. As same as in PCA, the singular value decomposition, which are core for dimensionality reduction techniques, is performed on the correspondence matrix P:

$$P = UDV^T, \tag{6}$$

where U and V are the orthogonal matrices representing singular vectors, which are actually coordinates in the reduced data space. The D is the diagonal matrix of containing singular values (square root of eigenvalues), which represent the inertia (a counterpart to variance in continuous variables) of an MCA, of each dimension. The total inertia of the dataset, and each singular value in D, represents the amount of inertia explained corresponding dimensions. Same as PCA, the cumulative inertia (scree plot) is commonly used to estimate the amount PCs retained. Lastly, we used an MCA biplot, by referring to clusters of PCA and binary factor variables, to come to conclusion which factors contribute the most to the reference value.

3. Results and discussion

Minitab 21® software was used for statistical analysis of experimental results. Each workpiece was measured according to the experiment plan and a total of 45 measurements were performed on one workpiece with 9 measurements at the central point. The adequacy of the model (R-sq) for all workpieces exceeds 90 %, which indicates that an adequate selection of factors was made and that the variations are well described by the obtained models. Analysis of variance (ANOVA) was used to determine statistically significant factors on the cylindricity of all workpieces. Fig. 3. proves the fact that the choice of measurement strategy can increase or decrease the measurement result by more than 100 %. The pictures show box-plot diagrams of the dispersion of the cylindricity measurement regarding different workpieces. In Fig. 4. p-values of individual factors are shown. It is possible to observe that the scanning speed and the stylus tip diameter are statistically significant for all workpieces since the p-value is below the significance threshold $\alpha = 0.05$ and a two-sided confidence interval of 95 %. The factor "number of points" is significant in the case of measuring cylindricity with the highest value of surface roughness (cylinder 11), while the filtration of sampled points is significant for cylinders 3 and 11. Also, a significant range of p-values for the number of points and filtration can be observed. Although the observation of the influence of individual factors leads to the conclusion that in most cases factors such as the number of points and the filtering of points should be omitted, the analysis of the interactions of individual factors leads to a different conclusion. For most workpieces, interactions of a number of points and filtration are significant.

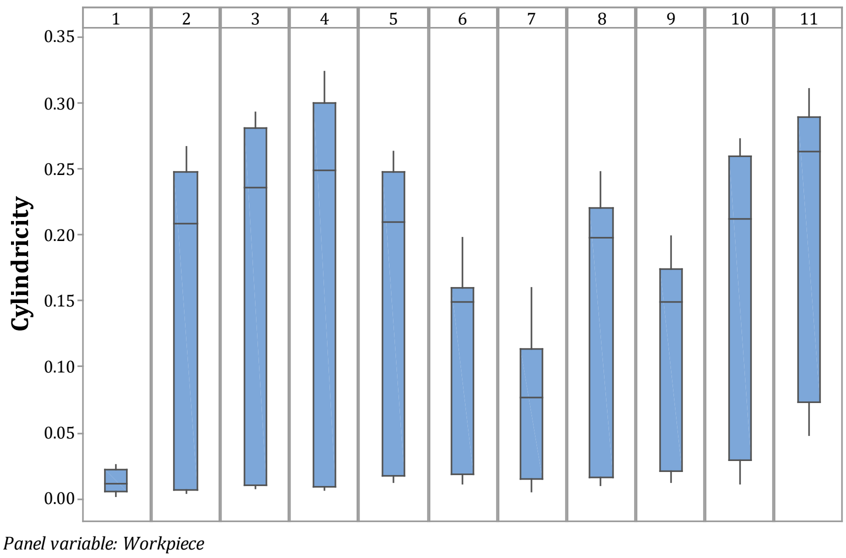
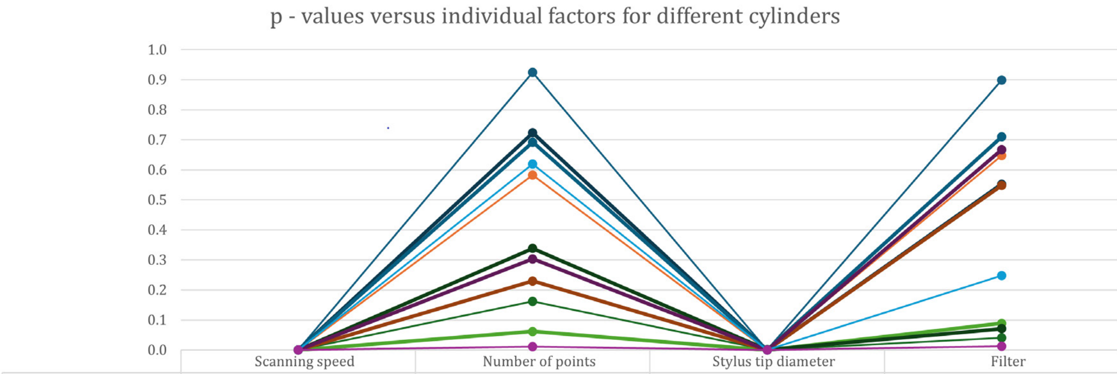


Fig. 3 Variability of measurement results depending on the applied measuring strategy

In Fig. 5 a combined diagram of the main effects for all workpieces is shown. The identification of optimal factor levels using the main effects diagram is usually a simple procedure where those factors that are closest to the optimal value defined by the user are selected. The optimal values, in this case, represent different values of cylindricity that were obtained on the roundtest (Table 1). As mentioned earlier, the workpieces differ significantly due to the different manufacturing processes. For example, in Fig. 6. reference values of cylindricity can be seen for cylinders 1, 5, and 11. For cylinder 1, for example, the optimal levels of the factors are: scanning speed – 5 mm/s; the number of points – 75; stylus tip diameter – 2; filter – cut off 8. For cylinder 5, the optimal factor levels are: scanning speed – 5 mm/s; the number of points – 25; stylus tip diameter – 5; filter – cut off 8. For cylinder 11, the optimal levels of factors are: scanning speed – 5 mm/s; the number of points – 25; stylus tip diameter – 5; filter - cut off 8. Identification of an optimal combination of factor levels for all workpieces is a serious challenge. When choosing a measurement strategy, the CMM programmer, in most cases, will not know the values of the quality characteristics such as the form error and the quality of the machined surface, and this way of finding the optimal values of the factors requires significant resources.



	Scanning speed	Number of points	Stylus tip diameter	Filter
p-values Cylinder 1	0.000	0.925	0.000	0.899
p-values Cylinder 2	0.000	0.582	0.000	0.648
p-values Cylinder 3	0.000	0.162	0.000	0.041
p-values Cylinder 4	0.000	0.062	0.000	0.089
p-values Cylinder 5	0.000	0.723	0.000	0.552
p-values Cylinder 6	0.000	0.230	0.000	0.548
p-values Cylinder 7	0.000	0.338	0.001	0.071
p-values Cylinder 8	0.000	0.692	0.000	0.710
p-values Cylinder 9	0.000	0.303	0.000	0.667
p-values Cylinder 10	0.000	0.619	0.000	0.248
p-values Cylinder 11	0.000	0.011	0.000	0.012

Fig. 4 *p*-values of individual factors for different cylinders

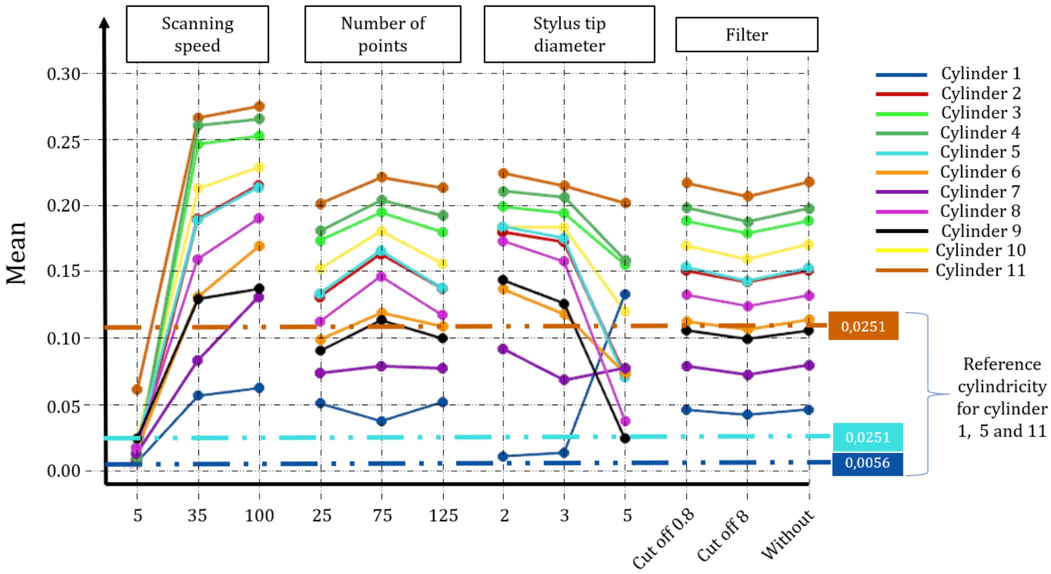


Fig. 5 Combined main effects plot for different workpieces

This study showed that the choice of factor levels and their percentage in the total variability of the measurement results depends on the real characteristics of the workpiece, even though they have the same nominal characteristics. In practical measurements, the real geometry of the workpiece is unknown, i.e. it is not determined before measuring on the CMM. To estimate optimal solution in the selection of the measurement strategy for measuring the cylindricity, we rely on unsupervised learning and dimensionality reduction tools of PCA (Principal Component Analysis) for quantitative data, and MCA (Multiple Correspondence Analysis) for qualitative data.

3.1 Principal component analysis result

The PCA scree plot (Fig. 6) shows that the first two components explain up to 83.8 % of the variance, while the first three components show 98.9 % of variance, suggesting excellent ordination properties. Interpreting the results in combination with all three components suggest that much of the information is retained. Indeed, observing the distribution of the points on PC' axes, the distributions (Fig. 7) suggest highest frequency points in PC1 and PC2. In fact, there are potential clusters observed by frequencies all three components, while PC4-PC6 retains only 1.1 % of the variance.

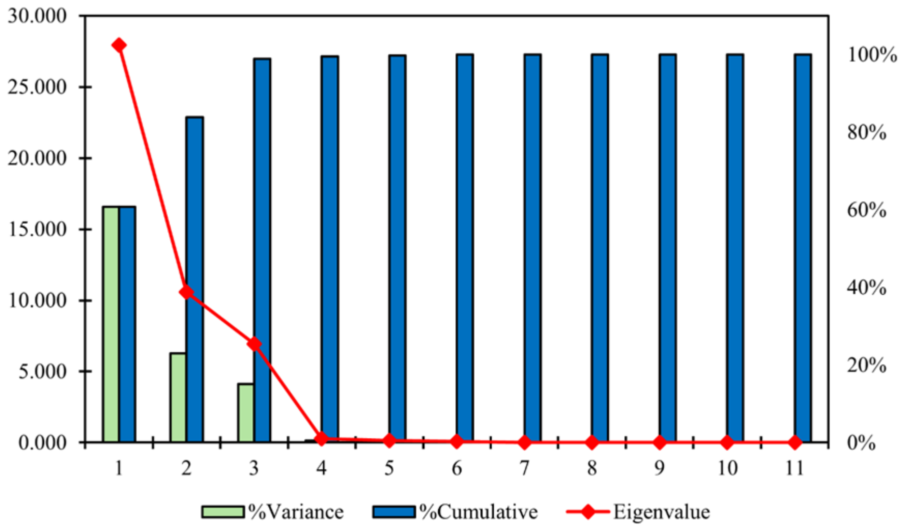


Fig. 6 Combined main effects plot for different workpieces

After projecting the points on the first three PCs, the results suggest three clusters (Fig. 8). In addition, we conducted additional exploratory analysis in the rest of the components, and the spread of the data did not suggest additional clustering, which is suggested by the distribution plots (Fig. 7). Given the reference value X46_Ref, our interest is to allocate clusters that share the most variance with this point. Thus, in all three ordinations these points are retained within the Cluster 1 (points X5, X10, X16, X20, X21, X23, X24, X27, X31, X32, X37, X44, X46).

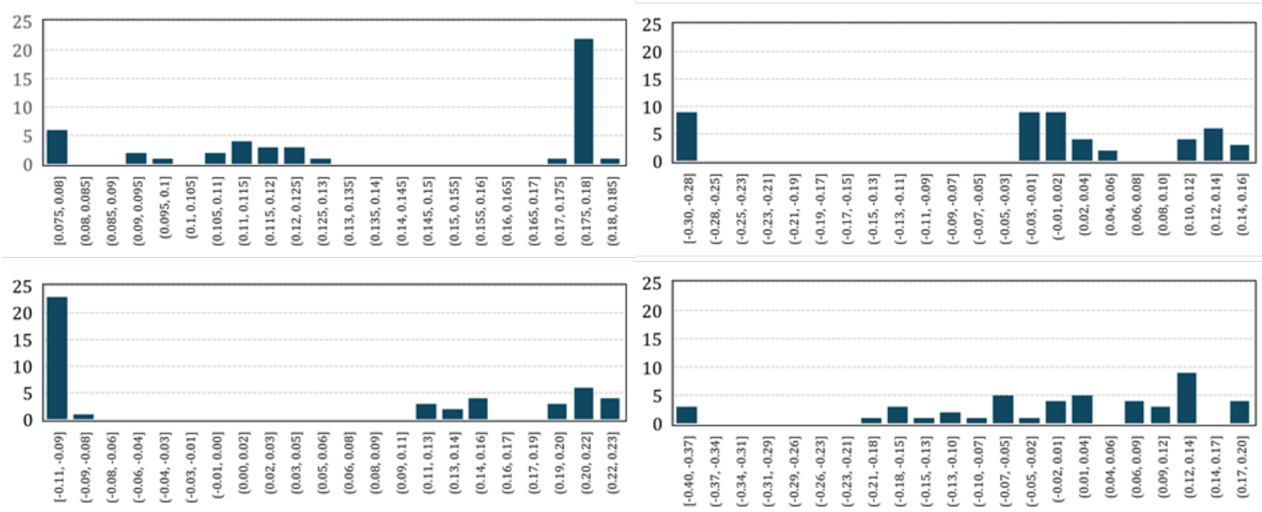


Fig. 7 Distribution plots of PC1 (top-left); PC2 (bottom-left); PC3 (top-right); PC4 (bottom-left)

After performing the PCA, labelling of clusters is performed (Fig. 9). Measurement speed 5 was perfectly associated with Cluster 1 (Fig. 9b). The results of the filter (Fig. 9a), and number of scanning points (Fig. 9d) did not provide good clustering properties, while using different measuring diameters (Fig. 9c), the results show almost perfect classification to Cluster 2. Nevertheless, given that performed clustering using factors from experimental setup did not provide full understanding of factors associating with reference value, we performed an MCA. The idea of performing an MCA analysis is to allocate factors that have the highest contribution to the associated measurements with respect to the reference value of measurements.

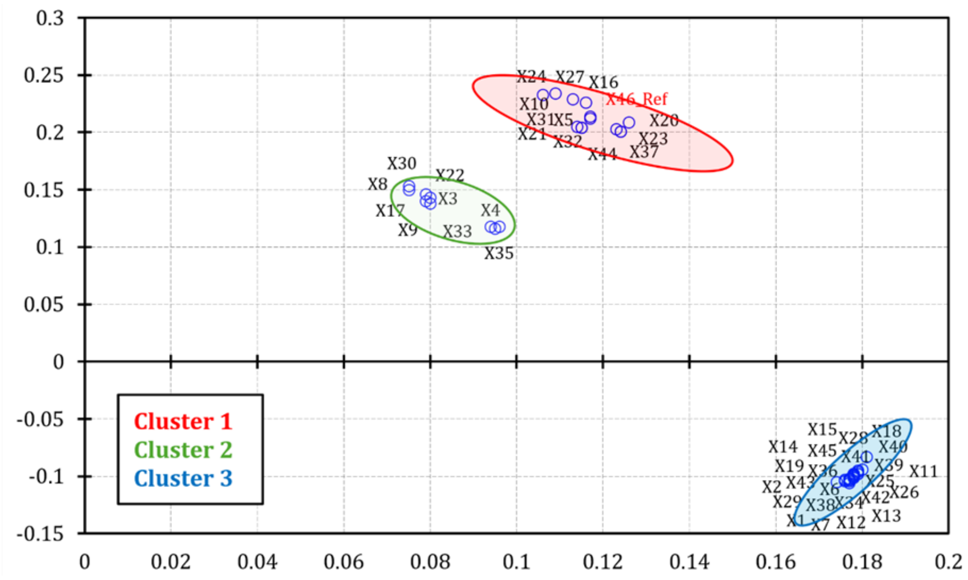


Fig. 8 Ordination using PC1 (x-axis) and PC2 (y-axis)

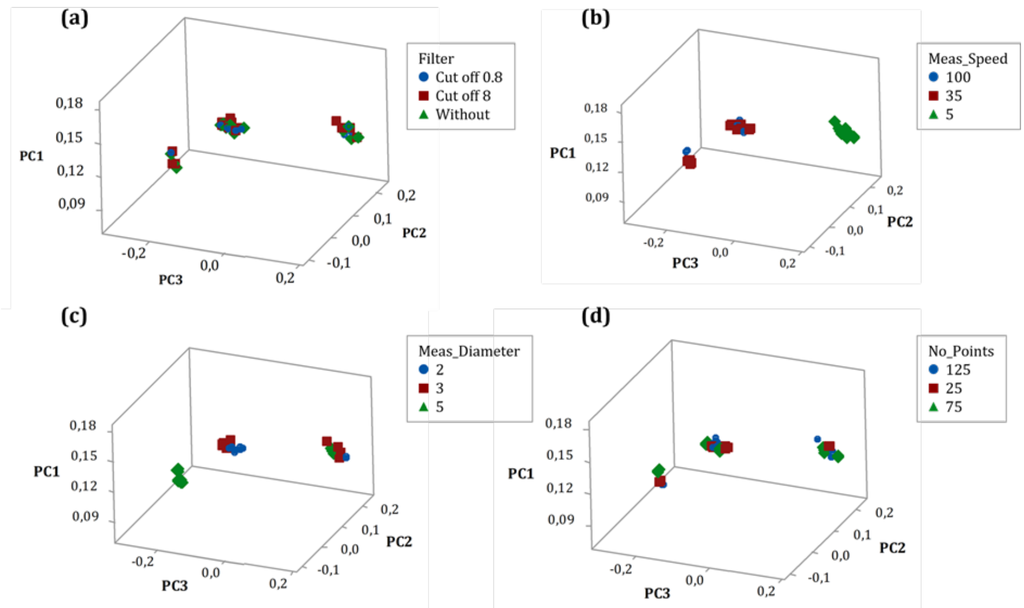


Fig. 9 PCA 3D ordination using classes (a) filter; (b) measurement speed; (c) diameter; (d) number of scanning points

3.2 Multiple correspondence analysis results

After performing a sequential construction of axes with the MCA, the eigenvalue decomposition shows that the first two components capture 38.75 % of the inertia (Fig. 10). Given that our main intent of using an MCA is to allocate factors associated with the Cluster 1, which contains the reference value, we relied on the quality of the interpretation and contribution associated with this cluster (Table 3). The MCA shows > 90 % quality of interpretation of clusters, where measurement speed 5 is perfectly association with Cluster 1, suggesting that measurement speed 5 has the highest impact for aligning the results of measurement to the reference value.

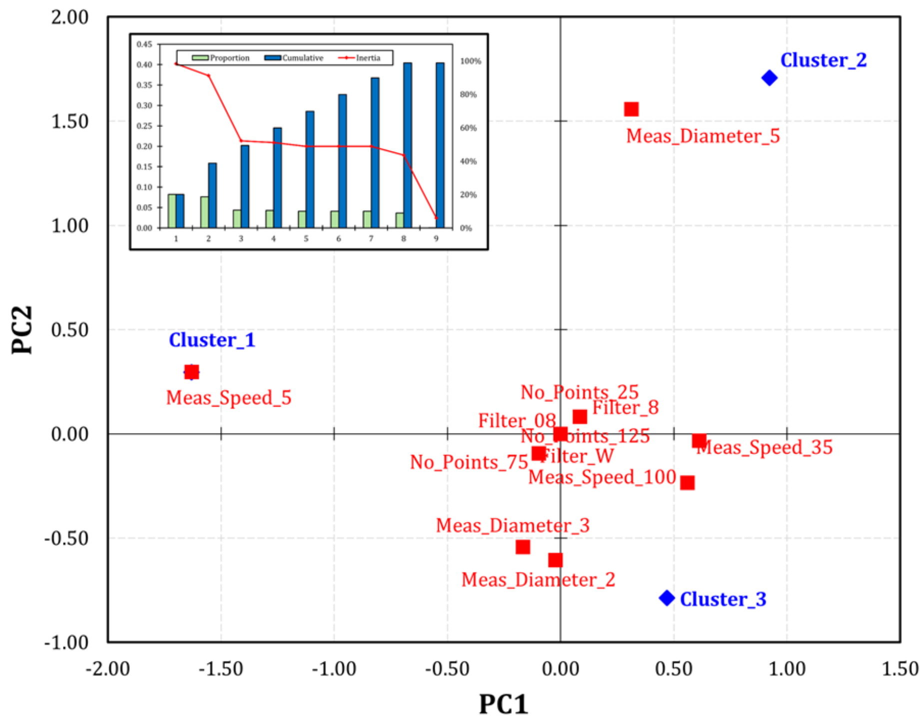


Fig. 10 MCA ordination biplot using PC1 (x-axis) and PC2 (y-axis) including the scree plot (top-left)

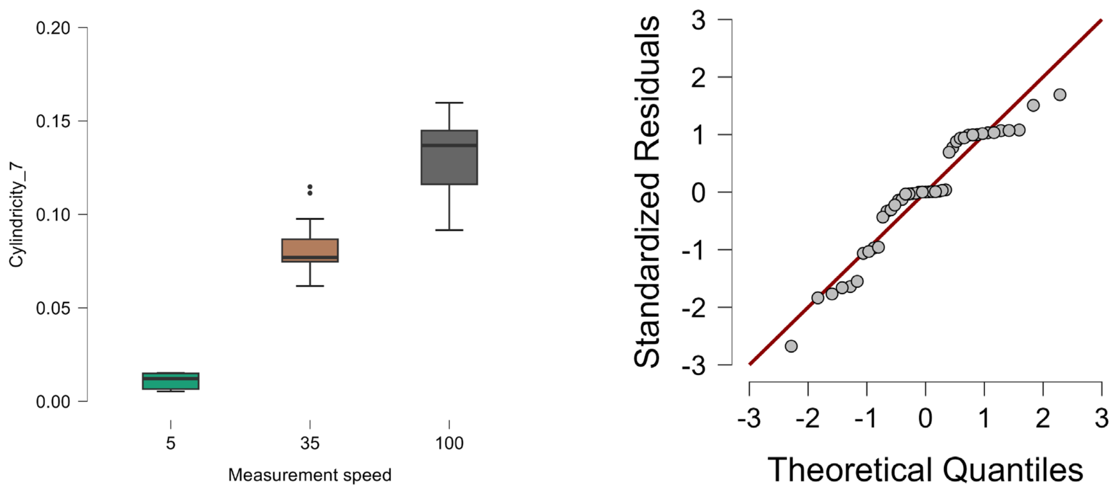
Table 3 Association between factors and clusters

Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
Cluster_1	0.998	0.053	0.073	-1.630	0.966	0.352	0.295	0.032	0.012
Cluster_2	0.942	0.040	0.080	0.923	0.213	0.085	1.708	0.729	0.313
Cluster_3	0.961	0.107	0.047	0.469	0.251	0.058	-0.788	0.710	0.178
Meas_Speed_5	0.998	0.053	0.073	-1.630	0.966	0.352	0.295	0.032	0.012
Meas_Speed_35	0.328	0.093	0.053	0.611	0.327	0.087	-0.034	0.001	0.000
Meas_Speed_100	0.134	0.053	0.073	0.560	0.114	0.042	-0.236	0.020	0.008
No_Points_25	0.005	0.053	0.073	0.085	0.003	0.001	0.082	0.002	0.001
No_Points_75	0.016	0.093	0.053	-0.097	0.008	0.002	-0.093	0.008	0.002
No_Points_125	0.005	0.053	0.073	0.085	0.003	0.001	0.082	0.002	0.001
Meas_Diameter_2	0.134	0.053	0.073	-0.023	0.000	0.000	-0.607	0.134	0.053
Meas_Diameter_3	0.282	0.093	0.053	-0.166	0.024	0.006	-0.542	0.257	0.074
Meas_Diameter_5	0.916	0.053	0.073	0.313	0.036	0.013	1.556	0.881	0.346
Filter_W	0.000	0.067	0.067	0.000	0.000	0.000	0.000	0.000	0.000
Filter_08	0.000	0.067	0.067	0.000	0.000	0.000	0.000	0.000	0.000
Filter_8	0.000	0.067	0.067	0.000	0.000	0.000	0.000	0.000	0.000

Note: Qual = Quality of interpretation; Mass = Proportion of the presence of the variable in complete dataset; Inert = Inertia; Corr = Correlation; Contr = Contribution to the cumulative inertia.

3.3 Independent Chi-square and ANOVA

To validate our findings we report χ^2 independence test. Namely, we used contingency tables formed between clusters and categorical variables. The obtained findings did not show any statistically significant dependency between clusters and filter. In fact, there was a spread of the same values across all categories ($\chi^2 = 0$; $p = 1.0$). The test statistic considering number of points also did not provide any statistically significant results ($\chi^2 = 0.804$; $p = 0.938$). However, the test considering measuring diameter reports statistically significant association with clusters ($\chi^2 = 33.549$; $p < 0.001$; VS-MPR = 28713.66). Most importantly, as observed in the MCA biplot (Fig. 11), the statistically significant relationship exists between clusters and measurement speed ($\chi^2 = 45.067$; $p < 0.001$; VS-MPR = $4.931 \cdot 10^6$), with reported high strength of the relationship (Cramer's $V = 0.708$). The performed analysis suggests that low measurement speed directly contributes to the improved cylindricity measurement "accuracy".



(a) (b)

Fig. 11 Box plot of measurement speed (a) and QQ plot (b)

Table 4 ANOVA Analysis

Cases	SSQ	df	MSQ	F	p	VS-MPR	η^2	η^2_r	ω^2
Measurement speed	0.589	2	0.294	225.971	< 0.01	$2.141 \cdot 10^{20}$	0.915	0.915	0.909
Residuals	0.055	42	0.001						

*SSQ = Sum of Squares; MSQ = Mean Sum of Squares; VS-MPR = Volk-Sellke Maximum p—Ratio (odds in favor of H_1 over H_0, estimated as 1/-e*p*log(p))*

Table 5 Bootstrapped Post Hoc Comparisons – Measurement Speed

Meas. speed	Mean	95%L	95%U	SE	*bias**	t	Cohen's d	p_{Tukey}	p_{Bonf}
5 : 35	-0.26	-0.276	-0.235	0.010	> 0.001	-19.665	-7.116	< 0.01	< 0.01
5 : 100	-0.26	-0.276	-0.245	0.008	> 0.001	-17.661	-7.251	< 0.01	< 0.01
35 : 100	-0.06	-0.029	0.023	0.013	> 0.001	-0.135	-0.135	0.927	1.00

****p < 0.001; * Bias corrected accelerated. Bootstrapping is based on 1000 succesful replicates. Mean difference estimate is based on the median of the bootstrap distribution. p value and confidence intervals adjusted for comapring family of 3 estimates (confidence intervals corrected using the Tukey method). The Bonferroni correction for multiple comparison test is also report as p_{Bonf}.*

Finally, relying on the results from the obtained cluster 1 (with measurement speed 5), we first subtracted values from the reference point within the cluster and use sum and average values with absolute deviation from the reference point (Fig. 12). What can be concluded from Fig. 12 is that low scanning speeds will give results closer to the reference value, while a uniform solution cannot be obtained for other factors. In general, it can be concluded that there are strong interactions between the number of points, the size of the diameter of the measuring probe, and filtering, and the operator should follow the guidelines from ISO 12180-2:2011 [27] regarding the Nyquist criterion when measuring.

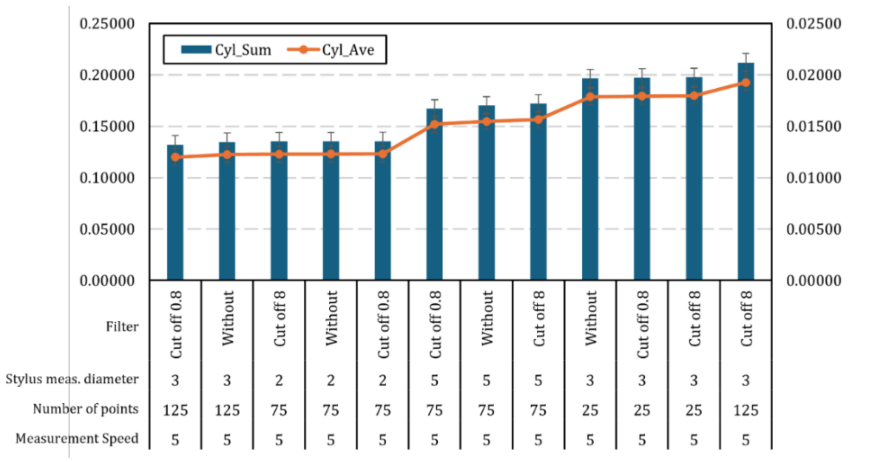


Fig. 12 Box plot of measurement speed (a), and QQ plot (b)

4. Conclusion

This paper deals with a practical industrial problem that CMM operators face every day, i.e. choosing a measurement strategy when measuring form error. A "bad" measurement strategy can lead to fatal decisions about rejecting good workpieces or letting bad workpieces pass the inspection. Since the complete specification operator is generally not indicated on the technical documentation, it is left to the machine operators to choose a strategy based on experience, while keeping in mind the economic imperative in terms of measurement time. Also in this study, the results of the measurements were compared with the reference value, and in this way, it was possible to reach the optimal levels of the considered factors. The general conclusion would be that when measuring cylindricity on a CMM, the scanning speed should be low and the choice of other factors must consider the interaction of the tip diameter of the measuring probe, the number of measuring points and filtering with the real geometry of the workpiece. As directions for future analysis, the research could be extended to other geometric deviations such as straightness, flatness, and circularity. A dependency could also be established when choosing the tip of the measuring probe and the quality of the machined surface of the workpieces, it can be determined how many points need to be measured and which filter to use.

Acknowledgement

This research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-65/2024-03/200156) and the Faculty of Technical Sciences, University of Novi Sad through project "Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad" (No. 01-3394/1).

References

[1] Jotić, G., Štrbac, B., Toth, T., Blanuša, V., Dovica, M., Hadžistević, M. (2023). The analysis of metrological characteristics of different coordinate measuring systems, *Tehnički Vjesnik – Technical Gazette,* Vol. 30, No. 1, 32-38, doi: 10.17559/TV-20220204091212.

[2] Dube, L., Gupta, K. (2023). Lean manufacturing based space utilization and motion waste reduction for efficiency enhancement in a machining shop: A case study, *Applied Engineering Letters*, Vol. 8, No. 3, 121-130, doi: 10.18485/aeletters.2023.8.3.4.

[3] Grazion, S., Spiryagin, V., Erofeev, M., Kravchenko, I., Kuznetsov, Y., Mukomela, M., Velichko, S., Ašonja, A., Kalashnikova, L. (2022). Diagnostics of defect detection in the initial stages of structural failure using the acoustic emission method of control, *Applied Engineering Letters*, Vol. 7, No. 2, 45-53, doi: 10.18485/aeletters.2022.7.2.1.

[4] Shen, Y., Ren, J., Huang, N., Zhang, Y., Zhang, X., Zhu, L. (2023). Surface form inspection with contact coordinate measurement: A review, *International Journal of Extreme Manufacturing*, Vol. 5, No. 2, Article No. 022006, doi: 10.1088/2631-7990/acc76e.

[5] JCGM 100:2008 (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement,* First edition, International Organization of Legal Metrology, Paris, France.

[6] ISO/TS15530-3:2011. *Geometrical products pecifications (GPS) – Coordinate measuring machines (CMM): technique for determining the uncertainty of measurement,* Part 3: Use of calibrated workpieces or standards, Geneva, Switzerland.

[7] ISO/TS15530-4:2008. *Geometrical product specifications (GPS) – Coordinate measuring machines (CMM): technique for determining the uncertainty of measurement,* Part 4: Evaluating task-specific measurement uncertainty using simulation, Geneva, Switzerland.

[8] Aggogeri, F., Barbato, G., Barini, E.M., Genta, G., Levi, R. (2011). Measurement uncertainty assessment of coordinate measuring machines by simulation and planned experimentation, *CIRP Journal of Manufacturing Science and Technology*, Vol. 4, No. 1, 51-56, doi: 10.1016/j.cirpj.2011.01.007.

[9] 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,* Vol. 158, Article No. 107692, doi: 10.1016/j.measurement.2020.107692.

[10] Acko, B., Brezovnik, S., Crepinsek Lipus, L., Klobucar, R. (2015). Verification of statistical calculations in interlaboratory comparisons by simulating input datasets, *International Journal of Simulation Modelling*, Vol. 14, No. 2, 227-237, doi: 10.2507/IJSIMM14(2)4.288.

[11] Lipus, L.C., Acko, B., Tompa, J. (2022). Experimental determination of influences on a gauge block's stack length, *Advances in Production Engineering & Management*, Vol. 17, No. 3, 339-349, doi: 10.14743/apem2022.3.440.

[12] Ricci, F., Scott, P.J., Jiang, X. (2013). A categorical model for uncertainty and cost management within the Geometrical Product Specification (GPS) framework, *Precision Engineering*, Vol. 37, No. 2, 265-274, doi: 10.1016/j.precisioneng.2012.09.005.
[13] Weckenmann, A., Knauer, M., Kunzmann, H. (1998). The influence of measurement strategy on the uncertainty of CMM-measurements, *CIRP Annals*, Vol. 47, No. 1, 451-454, doi: 10.1016/S0007-8506(07)62872-8.
[14] Barini, E.M., Tosello, G., De Chiffre, L. (2010). Uncertainty analysis of point-by-point sampling complex surfaces using touch probe CMMs: DOE for complex surfaces verification with CMM, *Precision Engineering*, Vol. 34, No. 1, 16-21, doi: 10.1016/j.precisioneng.2009.06.009.
[15] Raghunandan, R., Venkateswara Rao, P. (2008). Selection of sampling points for accurate evaluation of flatness error using coordinate measuring machine, *Journal of Materials Processing Technology*, Vol. 202, No. 1-3, 240-245, doi: 10.1016/j.jmatprotec.2007.09.066.
[16] 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*, Vol. 21, No. 6, 158-167, doi: 10.2478/msr-2021-0022.
[17] Orošnjak, M., Šević, D. (2023). Benchmarking maintenance practices for allocating features affecting hydraulic system maintenance: A West-Balkan perspective, *Mathematics*, Vol. 11, No. 18, Article No. 3816, doi: 10.3390/math11183816.
[18] Özel, T. (2006). Precision tracking control of a horizontal arm coordinate measuring machine in the presence of dynamic flexibilities, *The International Journal of Advanced Manufacturing Technology*, Vol. 27, 960-968, doi: 10.1007/s00170-004-2292-3.
[19] Stępień, K. (2015). An analysis of influence of sampling strategy and scanning speed on estimation of straightness and flatness deviations with CMMs, *Advanced Technologies in Mechanics*, Vol. 2, No. 2-3, 2-7.
[20] Ali, S.H.R. (2014). Performance investigation of CMM measurement quality using flick standard, *Journal of Quality and Reliability Engineering,* Vol. 2014, No. 1, Article No. 960649, doi: 10.1155/2014/960649.
[21] Urban, J., Beranek, L., Koptiš, M., Šimota, J., Košťák, O. (2020). Influence of CMM scanning speed and inspected feature size on an accuracy of size and form measurement, *Manufacturing Technology*, Vol. 20, No. 4, 538-544, doi: 10.21062/mft.2020.074.
[22] Magdziak, M. (2020). Determining the strategy of contact measurements based on results of non-contact coordinate measurements, *Procedia Manufacturing*, Vol. 51, 337-344, doi: 10.1016/j.promfg.2020.10.048.
[23] Rajamohan, G., Shunmugam, M.S., Samuel, G.L. (2011). Effect of probe size and measurement strategies on assessment of freeform profile deviations using coordinate measuring machine, *Measurement*, Vol. 44, No. 5, 832-841, doi: 10.1016/j.measurement.2011.01.020.
[24] Malburg, M.C. (2002). Fitting, filtering and analysis: Feature extraction in dimensional metrology applications, *International Dimensional Workshop*, *https://digitalmetrology.com/Papers/IDW2002-Slides.pdf*
[25] Lou, S., Brown, S.B., Sun, W., Zeng, W., Jiang, X., Scott, P.J. (2019). An investigation of the mechanical filtering effect of tactile CMM in the measurement of additively manufactured parts, *Measurement*, Vol. 144, 173-182, doi: 10.1016/j.measurement.2019.04.066.
[26] Li, K., Li, D., Ma, H.Q. (2023). An improved discrete particle swarm optimization approach for a multi-objective optimization model of an urban logistics distribution network considering traffic congestion, *Advances in Production Engineering & Management,* Vol. 18, No. 2, 211-224, doi: 10.14743/apem2023.2.468.
[27] ISO 12180-2:2011, *Geometrical product specifications (GPS) – Cylindricity,* Part 2: Specification operators, Geneva, Switzerland.