Prognosis and Multiobjective Optimization of the Sampling Plan for Cylindricity Evaluation

1. Introduction

Coordinate measuring machine (CMM) has revolutionized the manufacturing industries, owing to its high accuracy and precision capabilities. It is being widely used in the automotive, aerospace, and medical industries to perform the part inspection. It is a complex machine, where different components and subcomponents with their varied performance influence the measurement outcome. Indeed, the CMM inspection process must be well defined and designed for successful and efficient performance [1]. With growing intricacy of parts and tighter tolerances, intelligible approaches and techniques are needed to effectively plan measurement on the CMM. According to Baldwin et al. [2] and Weckenmann et al. [3], numerous variables as depicted in Figure 1 can influence the output of the CMM measurement. These factors encompass the sampling strategy, evaluation algorithm, workpiece position and orientation, surface conditions, sensor type and configuration, environment conditions, etc.
As reported by Moroni and Petrò [5], the sampling scheme is one of the parameters that significantly influences the measurement outcome. The inspection output from the CMM comprises a set of digitized surface points [6]. Each of these surface inspection points carries valuable information about the geometry of the measurand [7]. Certainly, this information is exploited to derive specifics regarding the variation in different geometries on the part. Verily, the quality and the volume of this information rely upon the density and the allocation of inspection points. The sampling strategy can be associated with the decisions regarding the number and spreading of points on the measuring surface [8]. As reported by Colosimo et al. [9], the inspection strategy is composed of two essential functions: the estimation of the sample size and the distribution of inspection points. The probed data acquired using the CMM are utilized to generate the substitute geometry, and this geometry is assessed against an ideal model. It follows the deviation between the extracted geometry using the sampled points, and the ideal geometry measures the accuracy on the part [1]. It implies that the evaluation of the form error is also an indispensable condition because it resolves the part’s compliance with the design requirements [10]. The data captured by the CMM have to be processed using the appropriate point sampling strategy as well as analyzed through a relevant evaluation algorithm to corroborate the component’s conformance to the customer needs. Henceforth, the sampling strategy and the computation of the geometrical deviation (evaluation algorithm) represent the important phases (as shown in Figure 2) which result in the prospective form error [11, 12].

The sampling strategy and the evaluation algorithm have been identified as the crucial steps in the inspection planning because they contribute substantially to CMM outcomes. Although the inspection error goes down with the increment of measurement points, it also inflates the inspection time, which raises the measurement cost, and thus the production expense. Hence, conducting CMM analysis using the least possible number of sample points (or sample size) is of greatest significance. However, the fewer number of points mandates the appropriate allocation of sampling points and the implementation of a suitable evaluation algorithm [9]. In fact, the evaluation of the form error employing just enough data points has been the emphasis for both researchers and practitioners [13]. More efforts are desired to establish the sampling plan and select the suitable evaluation algorithm for the inspection of geometric features effectively and efficiently.

The recognition of an appropriate sampling strategy for the enhanced CMM performance is a major concern and an appealing task. Multiple approaches have been used to describe the sampling strategy, but there is practically no study undertaken to predict form error and inspection time using machine learning (ML) tools. Furthermore, an inclusive study involving the optimization of the inspection plan using grey relational analysis (GRA) and fuzzy technique for order performance by similarity to ideal solution (FTOPSIS) is inadequate in earlier works. The primary objective of this work is therefore to establish prediction models and to optimize the inspection plan (or sampling strategy) in the evaluation of form error, particularly cylindricity. There are three primary tasks in this study. The first one is the identification of predictive models that can provide superior predictive accuracy with a combination of the sampling strategy and the evaluation algorithm. Various ML regression methods are studied and analyzed on the experimental data in the MATLAB regression learner toolbox [14]. Subsequently, GRA is carried out to acquire a reliable and dependable measurement plan for cylindricity assessment, which defines the second objective. The estimated values using GRA have also been confirmed through the execution of FTOPSIS. Finally, the assessment models for measurement uncertainty are realized to further substantiate the inspection output.

The rest of the paper is set up as follows. In Section 2, the literature review related to the determination of the sampling strategy, evaluation algorithm, and measurement uncertainty has been presented. Section 3 elaborates on the different ML
prediction tools and optimization methods, while in Section 4, materials, methods, and experimental work are discussed. In Section 5, the results are presented as well as discussed, and then the work is concluded in Section 6.

2. Literature Survey

The huge sample size does not guarantee a precise assessment of the form error. Indeed, the uncertainty in the form error evaluation declines with an improvement in the sample size only up to a particular value [15]. Therefore, several approaches have been developed in the last few decades to accomplish the relevant sample size. For instance, the statistical techniques, namely, the standard deviation and hypothesis testing, as reported by Menq et al. [16], can be employed to estimate the suitable sample size. Similarly, the feature-based statistical method was established by Jiang and Chiu [17] to compute the suitable volume of inspection points in two-dimensional (2D) features. This approach needed an agreeable threshold error as the decision criterion as well as regression and least square (LSQ) techniques to validate sample size prerequisite for satisfactory requirements. Furthermore, the adaptive sampling method was introduced by Edgeworth and Wilhem [18] to determine the suitable number of inspection points. They coupled the sample size to the quality of the product, where the sample size was highly dependent on the quality. The fuzzy logic, which effectively addresses the problems with imprecise and incomplete information, has also been implemented by many authors. For example, Lee et al. [19], Cho et al. [20], and Cho et al. [21] elaborated a fuzzy system for identifying the optimum number of inspection points. The low discrepancy algorithm, namely, Hammersley’s algorithm, was also employed to locate the measuring points on the target surfaces. In addition to the numerical and fuzzy notions in the sample size quantification, meta-heuristic applications (including genetic algorithm (GA), particle swarm optimization (PSO), and simulated annealing (SA)) can also be acknowledged. For instance, the GA approach devised by Meo et al. [15] to select the required number of sample points can be pursued. Correspondingly, Moroni and Petrò [22] introduced an inspection cost model to optimize the sampling strategy depending on the abatement of the inspection cost. They addressed the impact of the sampling strategy on inspection and its inconsistency cost. The application of SA was employed to reduce the cost of inspection and acquire an efficient sampling policy. To obtain an acceptable sample size, a search-based sampling technique developed by Badar et al. [23] may also be applied. The region-elimination search process, the tabu search, and the hybrid search (coordinate search, Hooke–Jeeves pattern search, and tabu search) were implemented for form error evaluation in their framework. Moreover, robust tools or practices like the design of experiment (DOE) are often used to measure the sample size acceptable for any specific application. Feng et al. [24] recommended and enacted DOE to obtain the befitting number of inspection points in hole evaluation. A statistical method based on the regression model was proposed by Pedone and Romano [25] to minimize the sample size. This technique was proved better than the well-known sampling techniques, including random sampling and Latin hypercube sampling. There have also been many works related to the allocation of inspection points on the measuring surface. For example, the inspection point strategies, namely, uniform, curvature dependent, and hybrid, developed by Pahk et al. [26] demonstrate one of the several leading efforts in this domain. Besides, Vrba et al. [27] also recognized the three different sampling strategies, involving the circumferential section method, the generatrix method, and the helix path method. Likewise, the spiral sampling method introduced by Guan et al. [28] can be exploited for form error assessment. Significant efforts can be listed out in the literature that were sought to identify the appropriate sampling strategy. However, a wide gap corresponding to the application of ML regression algorithms in form error prediction can be readily noticed. The ML prediction algorithms have been prevalent in numerous fields, but their utilization is somehow deficient in the field metrology, especially in the minimization of the inspection cost.

The ML algorithms can be exploited to control the response of a system for predicting a real value [29]. A myriad of manufacturing processes or methods can be formulated as regression problems and assessed utilizing ML techniques to establish predictive models [30]. For instance, Li et al. [31] utilized the data-driven ML-based predictive modeling approach to evaluate the surface roughness of the parts in the fused filament fabrication process. They employed different algorithms, namely, random forests (RFs), AdaBoost, classification and regression trees (CART), support vector regression (SVR), ridge regression (RR), and random vector functional link (RVFL) network. Likewise, Tapia et al. [32] also constructed a Gaussian process regression model for the estimation of porosity in metallic components, fabricated through the selective laser melting (SLM) process. Besides, Batyrnska et al. [33] provided a theoretical basis to exploit ML methods in the optimization of the powder bed fusion (PBF) manufacturing process. Countless research studies
concerning the prediction of process variables in machining can also be pointed out in the literature. These works unfolded the potential of ML to the manufacturing or machining processes, such as milling and turning. Wu et al. [34] proposed an RFs-based forecasting approach for the prediction of tool wear in the milling operation. The predictive model developed by RFs was the most accurate when compared with artificial neural networks (ANNs) and SVR. Similarly, Gouarir et al. [35] devised a prediction system depending on the convolutional neural network (CNN) to estimate the tool wear in the milling process. Mohammadi et al. [36] utilized the SVR classification algorithm and developed a prognostic model to forecast quality in the abrasion-resistant production method. In analogous work, Wu et al. [37] unified RFs and MapReduce data processing schemes to forecast tool wear in milling experiments. With this new approach, they attained a significant enhancement in processing speed and excellent prediction accuracy. Furthermore, the ML based on the multilayer perceptron (MLP) technique was applied by Saadallah et al. [38] to model the milling process for analyzing and improving its performance. There have been innumerable statistical techniques, with varied performances, introduced by the data mining and ML society [39]. Indeed, the performance of various statistical models was analyzed by Loyer et al. [40]. They obtained the predictive models using gradient boosted trees (GBT) and SVR for determining the production cost of the jet engine. They reported that the performance of GBT and SVR was better than that of multiple linear regression (MLR) and ANNs. Similarly, Jurkovic et al. [41] examined the performances of SVR, polynomial (quadratic) regression, and ANN for predicting the machining parameters in the high-speed turning process. They noticed that polynomial regression surpassed SVR and ANN with relation to cutting force and surface roughness prediction, while ANN exhibited superior performance in the prediction of tool lifetime. Although an abundance of work has been undertaken to study ML applications in manufacturing methods previously, there is rarely any implementation for inspection planning as yet.

It is of considerable relevance to incorporate uncertainty values in the measurement results for accomplishing a reliable inspection plan. Numerous studies related to the evaluation of measurement uncertainty can be pointed out in the literature. Barini et al. [42] conducted an uncertainty analysis on the complex product through studying the impact of point density as well as their distribution, probing system, and alignment. Similarly, the effect of various variables was investigated by Aggogeri et al. [43] through structured experimentation and a simulative method using a bootstrap algorithm. They identified the inspection part as the most significant parameter. The statistical method based on analysis of variance (ANOVA) was implemented by Piratelli-Filho and Di Giacomo [44] to assess the CMM measurement uncertainty as well as the influence of different variables. Various probe setups were reviewed by Caiazzo et al. [45] to quantify the type-A probe qualification uncertainty. Additionally, the kinematic factors, comprising speed, approach distance, and measurement strategy factors, were also examined. The statistical technique based on the Levene test for homogeneity of variances was utilized to analyze the outcome. According to Weckenmann et al. [4], operators and workpiece shapes should also be considered as the critical factors because they contribute significantly to the measurement uncertainty. A simplified mathematical model was developed by Arencibia et al. [46] to compute measurement uncertainty relevant to the inspection of circularity and cylindricity deviations. They recognized the probing error as having the greatest impact on the uncertainty values. Likewise, the random geometric error of the probing system was characterized by Weckenmann et al. [47] through an analytical approach.

While earlier research efforts have demonstrated the potentials of ML in different manufacturing applications, hardly any research can be reported on predicting inspection sampling plans. To address this research void, robust prognostic models have been derived by comparing different ML regression approaches. Consequently, the optimization techniques based on GRA and FTOPSIS are applied to obtain optimal parameters for the evaluation of cylindricity and inspection time. The uncertainty associated with the measurement has also been reported to establish a sense of confidence in the computed results.

3. Overview of Practiced Methods

This section presents the main characteristics and a summary of different ML regression techniques utilized in predictive modeling. It also outlines the multiobjective optimization methods employed to achieve the appropriate sampling plan.

3.1. ML Algorithms. ML can be described as a tract in artificial intelligence (AI) [48]. The architect of AI, Arthur Samuel, who contrived the term ML, explained it as a means of programming, which provides the computer the intelligence to learn [49]. It can be identified as one of the fundamental pillars or key enablers in smart manufacturing or Industry 4.0. It offers intuitiveness and understanding regarding complex processes without needing extensive expert insights [50]. Typically, ML involves a set of inputs or parameters to predict the outcome of a process in advance. Contrary to quintessential simulation procedures, where loads of work is needed to build up a model, ML algorithms acquire information or learn from the training data. ML can be classified depending on two primary learning techniques, particularly supervised and unsupervised learning [51]. The supervised learning sets up and trains a model using a set of established inputs and known output information for predicting future outcomes or responses to new data. The different supervised ML regression algorithms can be recognized as linear or nonlinear regression models, decision trees, neural networks, SVR, etc. On the other hand, the unsupervised learning utilizes only the input data and determines unexplored patterns or interrelations in the input
data. The general unsupervised ML algorithms include hidden Markov models, K-means, and Gaussian mixture models. Consequently, the outline and mathematical model of the ML regression approaches employed in this work are demonstrated.

3.1.1. Linear Regression. Linear regression (LR) can be explained as a mathematical model that demonstrates relationship between a dependent or a continuous response variable and one or more independent or predictor variables [29]. It can be described as a monotonic transformation of the linear combination of variables [39, 52]. The generalized form of an LR model can be represented using the following equation:

\[ y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \epsilon, \]

where \( y \) depicts a response or a dependent variable, \( x_i (x_1, x_2, x_3, \ldots, x_n) \) represents predictors or independent variables, \( \beta \) symbolizes linear parameter estimates or weights for each input attribute (\( \beta_0, \beta_1, \beta_2, \ldots, \beta_n \)), and \( \epsilon \) means the error terms. The weights have to be chosen to minimize the sum of squared differences between the actual and the predicted output values as in equation (2). The primary benefit of a linear model is its competence to interpret the significance of variables through coefficients as weights. However, it is susceptible to overfitting the interacting variables. The least-squares regression is carried out to establish linear relations in the training data:

\[ \sum_{j=1}^{k} \left( x_j^i - \sum_{i=0}^{n} \beta_i x_j^i \right)^2. \]

3.1.2. Decision Tree. The decision tree can be described as the predictive modeling approach that is based on an algorithmic approach to determine the means of splitting a set of data depending on various situations [53]. It is logically constructed through bifurcating the defined set of events along with a variable that minimizes the variance following the split [54]. The limitation of a decision tree is that it tends to overfit deep down in the tree since the decisions are dependent on fewer events. A decision tree is built through recursive partitioning—beginning from the seed node (known as the first parent), each node can be divided into subsequent child nodes. These nodes can be split further and become parent nodes themselves for their emanating children nodes. An example of a decision tree in Figure 3 forecasts the requirement for a machine maintenance.

The decision regression trees require the response variable to take continuous values (generally real numbers). The decision tree model is exceptional in managing tabular data with numerical or categorical attributes involving less than hundreds of categories [55]. In contrary to linear models, decision trees can analyze non-linear interaction between the input variables and the response value. The decision regression tree algorithm employs the concept of standard deviation reduction (SDR) to construct the decision tree [56]. Initially, the standard deviation of the dependent variable is estimated. It is followed by the splitting of datasets on various input variables and the computation of SDR of each branch before splitting. The attribute with the highest value of SDR is designated as the decision node. The dataset is further subdivided depending on the selected attributes, and if the standard deviation of a branch is greater than zero, it is split again. The process keeps on repeating until all the data are processed.

3.1.3. Support Vector Regression. Support vector machine (SVM) analysis, pioneered by Vladimir Vapnik, has been a prominent ML tool for classification and regression [57]. It is a nonparametric, kernel-based learning algorithm that translates the input information to a higher-dimensional feature space through nonlinear mapping to resolve an LR issue in this feature space. This transformation can be accomplished through either kernel functions, namely, linear kernel, polynomial kernel, Gaussian kernel, or sigmoid kernel [58]. The functioning of SVM is based on the concept of determining a hyperplane that is located at the highest distance from the two sets of points on either side of the plane [59, 60]. Its objective is to find the weights that maximize the bound. The points on the two sides of the plane are defined as the support vectors (SV), and they are the points that exhibit the greatest effect on the location of the plane. The SVM employed for regression problems is known as SVR [41]. Statistics and Machine Learning Toolbox™ employed in this work is based on linear epsilon-insensitive SVR (\( \epsilon \)-SVM), which is also known as L1 loss [14]. In \( \epsilon \)-SVR, the purpose is to estimate a function \( f(x) \) that possesses just \( \epsilon \) deviation from the acquired target surface, defined by the training dataset, and concurrently, it should be as flat as possible, as reported in [57]. The formulation for a separating plane in the input space can be defined using the following equation:

\[ f(x) = \beta^T x + b, \]

where \( \beta \) represents the weight vector normal to the plane, \( \beta^T x \) is the dot product in the input space \( X \in \mathbb{R}^n \), and \( b \in \mathbb{R} \) serves as bias, i.e., the distance of the hyperplane from the origin. The flatness of function \( f(x) \) can be accomplished through small weights, so the objective turns into the problem of minimizing the weight vector. \( \epsilon \)-SVM can be defined mathematically as an optimization problem using equation (4) [61]. Consider a training set \( \{(x_1, y_1), \ldots, (x_n, y_n)\} \)
and $C > 0$, $\varepsilon > 0$, where $x_i \in \mathbb{R}^m$ is a feature space and $y_i \in \mathbb{R}^1$ is the response output:

$$
\min_{a, a^*} \frac{1}{2} (a_i - a_i^*)^T Q (a - a^*) + \varepsilon \sum_{i=1}^l (a_i - a_i^*)
+ \sum_{i=1}^l y_i (a_i - a_i^*),
$$

subjected to $0 \leq a_i, a_i^* \leq C, i = 1, \ldots, l$, $\sum_{i=1}^l (a_i - a_i^*) = 0$, (4)

where $C$ is the tuning cost upper bound, $\varepsilon$ is a regularization parameter called the insensitive loss function, $Q$ is an $l \times l$ positive semidefinite matrix $Q_{ij} = y_i y_j K \cdot (x_i, x_j)$, and $K (x_i, x_j) = \phi (x_i)^T \phi (x_j)$ represents the radial basis function kernel (RBF) $e^{-\|x_i - x_j\|^2}$.

The SVR is very useful to abstain from the complexities of linear functions in the high-dimensional feature space, by evolving the problem of optimization into dual convex quadratic programs [62]. In SVR, the loss function is employed to correct accuracies that exceed the threshold $\varepsilon$. Such loss functions result in the sparse representation of the decision rule, providing significant algorithmic and symbolic benefits. The SVR, however, does not provide descriptive insights, such as probability estimates into the outcomes of the model; indeed, they are estimated utilizing uneconomical, fivefold cross-validation.

3.1.4. Ensemble Methods. The ensemble methods are powerful ML algorithms, which unified the multiple models [63, 64]. The ensemble learning algorithm strengthens weak learners and enhances the ability to predict. They are very stable, robust to outliers, and can handle nonlinearity efficiently. However, they are very complex, requiring higher computing power and resources. Bagging, also known as bootstrap aggregation, regularly extracts split datasets from the entire training dataset [29, 34, 65]. The conclusive predictions are estimated by averaging the results from all the decision trees generated on bootstrapped training subsets. Its primary objective is the minimization of the variability of a decision tree. It involves the formation of several subsamples of data from the training dataset, drawn at random with substitution. Each of these subsets is utilized to train the decision trees. Finally, the prediction is made by averaging all the predictions from individual trees. In contrast to bagging, the boosting approach creates the trees sequentially, where each subsequent tree is fit on the residuals of the previous model [66]. The objective is to reduce the error from the earlier tree at every step. Since the bagging approach is premised on mean predictions from subset trees, therefore, it may not always provide exact estimates for the regression model. Similarly, the boosting approach is prone to overfitting and needs a judicious tuning of different hyperparameters.

3.2. Multiobjective Optimization. Multiobjective optimization is imperative to control the process parameters without undermining any response parameter [67, 68]. It is indispensable to address the complexity of contemporary processes and accomplish more dependable decisions. However, it can be exhausting and demanding because the improvement or degradation of one performance characteristic may influence the other output attributes. Lately, numerous techniques have been used to work out multiresponse problems depending upon the case and intended accuracy. In this work, GRA and FTOPSIS have been utilized to optimize the multiobjective sampling plan. The grey hypothesis and fuzzy-based multiattribute decision-making (MCDM) methods are capable of capturing the incoherence and uncertainty in judgments and data gathering. Prof. Deng was the first who initiated the concept of grey theory from the grey set in consolidation with theories of space and control [69]. Likewise, the endeavor of Bellman and Zadeh [70] may be viewed as among the first extensions of fuzzy sets within MCDM. The FTOPSIS is one of the renowned fuzzy MCDM approaches that can be effectively employed in multiobjective optimization [71–73]. Amongst the most critical considerations in FTOPSIS is to measure the distance between two fuzzy numbers [74]. A straightforward and reliable vertex method introduced by Chen
[71] can be utilized to compute the distance between two triangular (or trapezoidal) fuzzy numbers. For example, if \( \bar{x} = (a_1, b_1, c_1) \) and \( \bar{y} = (a_2, b_2, c_2) \) symbolize the two triangular fuzzy numbers, the distance between these two fuzzy numbers is equal to

\[
\left( \frac{1}{4} \right) \left[ \left( a_1 - a_2 \right)^2 + \left( b_1 - b_2 \right)^2 + \left( c_1 - c_2 \right)^2 \right].
\] (5)

If \( \bar{p} = (a_1, b_1, c_1, d_1) \) and \( \bar{q} = (a_2, b_2, c_2, d_2) \) represent the two trapezoidal fuzzy numbers, then the distance is as follows:

\[
f(\bar{p}, \bar{q}) = \left( \frac{1}{4} \right) \left[ \left( a_1 - a_2 \right)^2 + \left( b_1 - b_2 \right)^2 + \left( c_1 - c_2 \right)^2 + \left( d_1 - d_2 \right)^2 \right].
\] (6)

The adoption of a satisfactory sampling technique in CMM inspection is a challenging task since there are multiple uncontrollable factors at play. These factors engender doubt, confusion, and uncertainty. They also render the definition of a purposeful sampling strategy a daunting activity. The metrologists usually ascertain the sampling model based on intuition or experience, which is not a feasible solution if a trade-off between inspection correctness and measurement speed is needed. The execution of GRA and FTOPSIS is henceforth particularly interesting in this work to generate decision-making or optimization models capable of handling insufficient or unclear knowledge and information inherent in the inspection process.

4. Proposed Methodology

This section discusses the procedure adopted to acquire the prediction models and realize the suitable sampling plan. It presents the experimental setup, and the workpiece is utilized to collect the data needed for subsequent analysis.

4.1. Specimen and Setup. A cylindrically shaped component is employed in this study to prepare the dataset. It is because cylindrical features are prevalent in design/manufacturing industries, and their correctness in assemblies is of significant relevance [46, 75]. The cylindrical test specimen as displayed in Figure 4(a) is selected for the experimentation. The error in cylindricity is the deviation of the estimated cylindrical surface and its ideal cylindrical counterpart [76]. This is a critical performance metric for the mechanical component’s precision and assembled items. The magnitude of cylindrical error seriously influences the wear rate, alignment, movement of rotating parts, and mechanical assembly effectiveness. The assessment of cylindrical error is much more complicated and nuanced compared with the computation of straightness, flatness, and circularity errors. The cylindricity error can be interpreted as the superposition of three types of deviations: a straightness variance in the cylinder axis, a shape deviation in the cylinder longitudinal plane, and a shape distortion in the cylinder cross section [77]. Precise cylindricity quantification, therefore, necessitates the application of sophisticated measurement equipment, such as the CMM.

The research utilizes the medium-sized CMM equipped with a touch-trigger probe. This CMM facilitates an accuracy of (1.6 + L/333) μm which follows DIN EN ISO 10360-2: 2001. This machine is indeed able to handle workpieces up to the size of 1200 mm/900 mm/700 mm. Also, the effective utilization of the CMM mandates the implementation of an optimal inspection plan, which contains an appropriate combination of the point distribution technique, the spacing of inspection points, the evaluation algorithm, etc.

4.2. Data Acquisition. A series of experiments premised on a full factorial design are undertaken with different combinations of factors defining the inspection plan. The variables, including the point distribution strategy, the distance between points, the number of point levels, and the evaluation algorithm, are considered in the inspection strategy [27]. The measurement accuracy and the time required are assigned as the performance indicators because they are the key determinants in ensuring the optimal inspection plan [78]. The guidelines for measuring the cylindricity are outlined in ISO 12180 standards. Following ISO 12180-2:2011 [79], various point allocation strategies, including roundness (circular-ferential) profiles, generatrix lines, bird cage, helix, and random points, can be employed [80]. This article has considered three approaches, namely, the roundness profiles, the generatrix lines, and the helix, as demonstrated in Figure 5. The roundness profiles and the generatrix lines strategies were adopted in most prior studies. Interestingly, the helix has not been widely used to determine variances in cylindricity although it can yield the most comprehensive data regarding the artifacts being gauged. In the generatrix method, the cylindrical surface is inspected in a sequence of straighter paths that are taken vertically along the cylinder axis, while the cylindricity is measured around the radial cross-sections in the roundness profile allocation technique. The roundness profiles distribution strategy derives surface data in a series of parallel circular planes, and the generatrix lines retrieve surface information in the form of line contours. The third sampling scheme is the helix pattern, which follows a centerline by constructing a helix of a fixed angle. Using these methods, especially roundness profiles and helix, adequate point density and detailed surface coverage can be assured. In the same way, various evaluation algorithms are used to compute the quality of a set of points for assessing the surfaces [77, 79]. In this research, four algorithms, namely, LSQ, minimum zone (MZ), maximum inscribed circle (MIC), and minimum circumscribed circle (MCC), are selected [81]. LSQ, MZ, MCC, and MIC are some of the most frequently utilized data fitting methods.
The implementation of these methods is based on the required applications. For example, MIC and MCC are required for applications involving mating parts. Until recently, the LSQ technique has been the most popular and widely used method. It is because LSQ is computationally simple and provides unique solutions. Although LSQ is very stable and less susceptible to local deviations and surface effects, it sometimes provides an approximate solution and leads to an overestimation of the error. The central principle of the MZ approach is the search for two coaxial cylindrical surfaces comprising all inspection points, and the radius difference between the two coaxial cylindrical surfaces is minimal. Because data interpretation is tedious in the MZ framework, several approximate and comparable precision mechanisms, namely, the MCC and MIC processes, have been found.
The remaining independent variable is the point density. The point density for different runs has been modified by varying the step width (point spacing) and the number of paths. The distribution strategies (roundness profiles, generatrix lines, and helix) describe the inspection points’ spread, while varying paths as well as step width results in variable sample sizes, as presented in Table 1. For example, a combination of 0.5 mm step width and 8 (the number of paths) results in a total of 1156 number of points in the circumferential point distribution. Similarly, the helix point distribution pattern with step width 1 mm and number of paths 8 provides a total number of points equal to 569.

The real cylindricity of the specimen has also been estimated to analyze the cylindricity error at various combinations of inspection parameters. The actual cylindricity is indeed exploited as the reference cylindricity for the successive experimental runs. The goal of the multiobjective optimization in this study is to accurately estimate the cylindricity of the specimen in lesser inspection time. Hence, the reference cylindricity has been derived by incorporating as many measuring points as possible, which can contribute to the largest achievable measuring accuracy. A great number of inspection points improve the likelihood of ascertaining the real cylindricity of the specimen. For this research, the reference cylindricity is measured using 10,000 points distributed over the length of the cylinder in roundness profiles. All the measurements are carried out in the scanning mode with the scanning speed of 2 mm/s, the probe sphere diameter of 3 mm, and the overall length of 50 mm. The CMM travelling speed is constant at 70 mm/s.

### Table 1: Factors and their corresponding levels.

| Distribution strategy | Evaluation algorithm | Step width (mm) | Number of points | Total number of points |
|-----------------------|----------------------|----------------|------------------|-----------------------|
| Roundness profiles    | LSQ                  | 0.5            | 2, 4, 6, 8       | 284, 568, 852, 1136   |
|                       | MZ                   | 1              | 2, 4, 6, 8       | 142, 284, 426, 568    |
|                       | MCC                  | 1.5            | 2, 4, 6, 8       | 96, 192, 288, 384     |
|                       | MIC                  | —              | —                | —                     |
| Generatrix lines      | LSQ                  | 0.5            | 3, 4, 6, 9       | 84, 112, 168, 252     |
|                       | MZ                   | 1              | 3, 4, 6, 9       | 42, 56, 84, 126       |
|                       | MCC                  | 1.5            | 3, 4, 6, 9       | 30, 40, 60, 90        |
|                       | MIC                  | —              | —                | —                     |
| Helix                 | LSQ                  | 0.5            | 2, 4, 6, 8       | 286, 569, 853, 1137   |
|                       | MZ                   | 1              | 2, 4, 6, 8       | 143, 283, 427, 569    |
|                       | MCC                  | 1.5            | 2, 4, 6, 8       | 96, 190, 285, 379     |
|                       | MIC                  | —              | —                | —                     |

4.3. Predictive Modeling. The primary objective of employing ML algorithms is the identification of an unexplored dependency between the input (or independent variables) and the output dependent variables in the dataset. In the current study, the regression techniques have been established as learning algorithms to anticipate form errors and inspection efficiency. The ML regression techniques based on LR, decision tree, SVR, and ensembles of trees are used in this study. These algorithms have been copiously utilized for modeling plentiful real-life regression problems. The motivation of using different ML algorithms is the pursuit of the most successful ML regression technique for the problem presented in this work. The dataset utilized in this investigation consists of output variables, namely, cylindricity and the measurement time, and four predictors that include both continuous (step widths and number of paths) as well as categorical variables (point distribution strategy and evaluation algorithm). The performance criteria used in this study to assess their prediction accuracy are root-mean-square error (RMSE), mean absolute error (MAE), and coefficient of determination ($R^2$). The RMSE, which quantifies the difference between the predicted and the true values, is computed using the average of the error squares [84]. The RMSE can be estimated using equation (7) [85]. The MAE, which measures the variability in the dataset, can be obtained through the average distance between all the estimated and the true values (see equation (8)). The $R^2$-squared measure assesses the spread of the predicted response around the fitted regression line:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (e_i - t_i)^2 + (e_2 - t_2)^2 + \cdots + (e_n - t_n)^2}{n}},$$

(7)

$$\text{MAE} = \frac{\sum_{i=1}^{n} |e_i - t_i|}{n},$$

(8)

where $e$ is the predicted value and $t$ represents the true value.

The RMSE and MAE are always positive, and their smaller values are needed for superior performance of ML algorithms. Furthermore, the $R^2$-squared measure is always within the band of 0 to 1, and its value near 1 is preferred. The performance indicators for these algorithms have been estimated utilizing the technique of fivefold cross-validation. In this approach, the entire initial dataset with 144 observations is arbitrarily split into five subsets of equal sizes. It means the training is conducted five times on four segregations, while the remaining solitary partition is employed to assess the regression model’s performance.
predictive effectiveness. Thus, the values of performance indices are averaged to estimate the model's predictive accuracy [86, 87].

4.4. Multiobjective Optimization. The multiobjective models have been developed to simultaneously minimize the cylindricity error and measurement time. As decision-making includes ambiguity, strategies such as GRA and FTOPSIS are reasonable options. These techniques are premised on viable mathematical approaches for system analysis characterized by incomplete information. The two approaches have been used to ascertain the results from each other. GRA and FTOPSIS are deployed to rate the various combinations of parameters and to establish the combination that can offer a trade-off between accuracy and time. GRA execution may be carried out as follows [88, 89].

Assume that \( s_j^*(l) \) denotes the string after the data processing, \( s_j^{(O)}(l) \) represents the original sequence of responses, (where \( j = 1, 2, \ldots, m \) and \( l = 1, 2, \ldots, n \)), \( \max s_j^{(O)}(l) \) means the largest value of \( s_j^{(O)}(l) \), and \( \min s_j^{(O)}(l) \) is the smallest value of \( s_j^{(O)}(l) \). The data have been standardized as follows:

\[
s_j^*(l) = \frac{s_j^{(O)}(l) - \min s_j^{(O)}(l)}{\max s_j^{(O)}(l) - \min s_j^{(O)}(l)} \text{ larger \ the \ better,}
\]

\[
s_j^*(l) = \frac{\max s_j^{(O)}(l) - s_j^{(O)}(l)}{\max s_j^{(O)}(l) - \min s_j^{(O)}(l)} \text{ smaller \ the \ better,}
\]

\[
s_j^*(l) = \frac{|s_j^{(O)}(l) - \text{IV}|}{\max \{\max s_j^{(O)}(l) - \text{IV}, \text{IV} - \min s_j^{(O)}(l)\}} \text{ nominal \ the \ better}
\]

\[
\cdot (\text{characterize the response as the ideal value (IV)).}
\]

A reference series is determined using the comparability sequence. The grey relational coefficient (GRC) is figured using the preprocessed sequence following data processing. The GRC can be estimated by applying the following equation:

\[
\text{GRC}(s_{o}^{*}(l), s_{j}^{*}(l)) = \frac{\Delta_{\min} + \xi \Delta_{\max}}{\Delta_{oj}(l) + \xi \Delta_{\max}}
\]

where \( \Delta_{oj}(l) \) characterizes the deviation sequence of the reference sequence \( s_{o}^{*}(l) \) and the comparability sequence \( s_{j}^{*}(l) \), i.e., \( \Delta_{oj}(l) = |s_{o}^{*}(l) - s_{j}^{*}(l)| \) represents the absolute value of the difference between \( s_{o}^{*}(l) \) and \( s_{j}^{*}(l) \):

\[
\Delta_{\min} = \min_{v} \min_{l} \Delta_{oj}(l),
\]

\[
\Delta_{\max} = \max_{v} \max_{l} \Delta_{oj}(l),
\]

where \( \xi \) is the distinguishing coefficient. GRA has just one parameter, called as a distinguishing coefficient (\( \xi \)), and its value lies between 0 and 1. This parameter is often needed to fine tune the range of the comparative setting and to regulate the extent of variations of the relational coefficients [90]. The distinguishing coefficient is intended to broaden or restrict the extent of the GRC [91]. Its assignment is dependent mostly on the judgment of the decision maker [92]. It is very important to estimate the value of the distinguishing coefficient, which is more reflective of the real situation (or true system uncertainties). Of this reason, FTOPSIS is also used in this study, and the distinguishing coefficient value in GRA is selected which provides the most comparable ranking to FTOPSIS. Kendall's concordance coefficient that has been discussed later in the section is applied to assess the ranking agreement between the two methods. Besides, the distinguished coefficient-guided sensitivity analysis is also implemented to strengthen the credibility of the acquired ranking order through the GRA model. Accordingly, if the ranking order is consistent for most of the changes in \( \xi \), this means that the ranking order is stable and effective.

The grey relational grade (GRG) can be described as a weighting sum of the GRC and is calculated utilizing the formulation below:

\[
\text{GRG}(s_{o}^{*}, s_{j}^{*}) = \sum_{l=1}^{n} \beta_{l}(s_{o}^{*}(l), s_{j}^{*}(l)),
\]

where \( \beta_{l} \) represents the weighting value of the \( l \)-th performance characteristic and \( \sum_{l=1}^{n} \beta_{l} = 1 \). Based on the application requirements, the selection of \( \beta \) for specific output characteristics should be chosen. For two performance attributes of equivalent significance, each of them can be given weights of 0.5. Finally, the rating for the various parameter combinations is dependent on the GRG.

The different steps in FTOPSIS using trapezoidal fuzzy numbers can be presented as follows [93, 94].

**Step 1.** Data representation by adopting the linguistic values.

The execution of an efficacious sampling strategy is a critical activity that incorporates many unknown variables and distorted data. The criterion of linguistic terms is therefore very helpful [69], particularly in the problem outlined in this study.

Due to their efficacy, clarity, ease of implementation, and general acceptance, the linguistic variables are characterized using trapezoidal fuzzy numbers in this paper. In other words, it can be stated that by leveraging trapezoidal
fuzzy numbers, the decision matrix (or the response matrix) has been represented. It is possible to specify the membership function of the trapezoidal fuzzy number labeled by $T$ and defined as $(a, b, c, d)$ using the following formulation [95]:

$$
\mu_{\tilde{T}}(x) = \begin{cases} 
0, & x \leq a, \\
\frac{x - a}{b - a}, & a \leq x \leq b, \\
1, & b \leq x \leq c, \\
\frac{d - x}{d - c}, & c \leq x \leq d, \\
0, & x \geq d.
\end{cases}
$$

(15)

The constants $(a, b, c, d)$ indicate the vagueness of the data being evaluated [96]. It implies the smaller the interval $[a, d]$, the lesser the fuzziness in the computing data.

**Step 2.** Estimation of a normalized fuzzy decision matrix.

The normalized fuzzy decision matrix can be visualized as $R = [\tilde{r}_{ij}]$. The responses (or the attributes) can be expressed either as a positive response (or benefit attribute) or a negative response (or cost attribute). The normalized fuzzy decision matrix can therefore be produced by employing the formulae below:

(i) **Positive Response.** Higher the value of the response, superior will be the performance:

$$
\tilde{r}_{ij} = \left\{ \frac{a_{ij}}{d_{ij}}, \frac{b_{ij}}{d_{ij}}, \frac{c_{ij}}{d_{ij}}, \frac{d_{ij}}{d_{ij}} \right\},
$$

(16)

$$
d_{ij}^+ = \max\{d_{ij}^j \}_{j \in \text{positive response}}.
$$

(ii) **Negative Response.** Lower the value of the response, better will be the output:

$$
\tilde{r}_{ij} = \left\{ \frac{a_{ij}}{d_{ij}}, \frac{a_{ij}^-}{c_{ij}^-}, \frac{a_{ij}^-}{b_{ij}^-}, \frac{a_{ij}^-}{a_{ij}} \right\},
$$

(17)

$$
a_{ij}^- = \min\{a_{ij}^j \}_{j \in \text{negative response}}.
$$

**Step 3.** Calculation of the weighted normalized decision matrix.

Depending on specifications, the weighted normalized decision matrix assigns significance to individual responses (or aspects). This matrix is determined by multiplying the normalized decision matrix with weights specified to the responses as follows:

$$
\tilde{M} = (\tilde{m}_{ij}), \quad \text{where} \quad \tilde{m}_{ij} = \tilde{r}_{ij} \ast w_j.
$$

(18)

**Step 4.** Estimation of the fuzzy positive ideal solution (FPIS) and the fuzzy negative ideal solution (FNIS). The FPIS and FNIS are computed as follows:

$$
A^+ = (\tilde{m}_{i1}^+, \tilde{m}_{i2}^+, \tilde{m}_{i3}^+, \ldots, \tilde{m}_{in}^+), \quad \text{where} \quad \tilde{m}_{ij}^+ = \max_i\{m_{ij}\},
$$

$$
A^- = (\tilde{m}_{i1}^-, \tilde{m}_{i2}^-, \tilde{m}_{i3}^-, \ldots, \tilde{m}_{in}^-), \quad \text{where} \quad \tilde{m}_{ij}^- = \min_i\{m_{ij}\}.
$$

(19)

**Step 5.** Determination of the distance from every option to the FPIS and the FNIS:

$$
f_{ij}^+ = \sum_{j=1}^{n} f(\tilde{m}_{ij}, \tilde{m}_{ij}^+),
$$

(20)

$$
f_{ij}^- = \sum_{j=1}^{n} f(\tilde{m}_{ij}, \tilde{m}_{ij}^-),
$$

where $f_{ij}^+$ and $f_{ij}^-$ symbolize the distance from each alternative to the FPIS and the FNIS, respectively.

**Step 6.** Calculation of the closeness coefficient (CC) for every alternative:

$$
CC_i = \frac{f_{ij}^-}{f_{ij}^+ + f_{ij}^-}.
$$

(21)

**Step 7.** Ranking the alternatives: the preferred alternative is that which has the greatest closeness coefficient.

Kendall’s concordance coefficient ($W$) [97, 98] can be estimated to quantify the degree of ranking consensus between different ranking methods. $W$ ranges from 0 to 1, with 1 implying the ideal compliance between the outcomes from various methods:

$$
W = \frac{12R}{m^2(k^2 - k)}.
$$

(22)

where $R = \sum_{i=1}^{m} r_{ij}$; $R = \sum_{i=1}^{k} (R_i - R)^2$, $r_{ij} =$ rating optimization method $i$ gives to combination $j$, $m =$ optimization methods, and $k =$ number of combinations.

4.5. Uncertainty Evaluation. There are precisely many other independent variables that might exist in the experiment besides four parameters used in this investigation. It is not possible to consider all the factors due to the high cost and effort of a large number of experiments. The remaining variables may consist of environmental conditions, workpiece properties, probing method, part fixtureing, etc. This is why the measurement uncertainty has been evaluated in this work. Measurement uncertainty as described in Figure 6 can be defined as the interval or range around the measurement result within which the true value of the measured quantity lies with some probability [46, 99]. It is hard to comprehend that perhaps the true measurement value can never be established, but the level of confidence in a measured value
can be conveyed by measurement uncertainty [100]. The parameters describing the spread of values such as standard deviation and half of the range having a certain probability can be used to estimate measurement uncertainty associated with a given measurand [101]. These parameters, which can be referred to as standard uncertainties, \( u_i \), are usually computed for several individual components. For instance, in CMM measurement, standard uncertainties related to the workpiece, machine errors, probing system, environment, etc., can be considered.

The uncertainty is evaluated through experiments by varying different uncertainty contributors. This approach is based on multiple measurement strategies and a non-calibrated workpiece. This approach has been inspired by the methodology developed within ISO/TS 15530-2:2004 [103]. The experimental approach is adopted because of its simplicity as well as its ability to provide realistic statements of measurement uncertainties. The different factors, including the number of measurement points, point distribution, evaluation algorithm, part orientation, and probe configuration, are considered. These factors, two at a time, are merged together to estimate the true value of the measurand. Indeed, the framework builds on the scientific approach that aims to develop a method for quantifying the measurement uncertainty.

The measurements are divided into three phases. Phase I comprises measurement of the test part using various combinations of the sampling strategy and the evaluation algorithm. This procedure quantifies measurement uncertainty due to point distribution, number of points, and evaluation algorithm. In phase II, CMM geometric errors, probing error, workpiece positioning error, and fixture errors are investigated through the measurement in several positions and orientations of a workpiece as well as the probing system in the work volume. Since, the workpiece shape can be estimated from the production strategy [4], therefore, to predict measurement uncertainty associated with workpiece characteristics, workpieces obtained under different machining conditions are considered in phase III. All the input quantities are assumed independent; that is, there is no correlation between them.

The measurement uncertainty due to the sampling strategy and evaluation algorithms involve statistical analysis (type-A uncertainty), which has been carried out using the following procedure.

Suppose \( n_s \) is the number of different point distributions \( (i = 1 \text{ to } 36) \), and \( n_t \) is the repetitions at each point distribution \( (k = 1 \text{ to } 2) \), and \( n_3 \) is the number of evaluation algorithms \( (j = 1 \text{ to } 4) \). The measurement uncertainty due to point distribution, \( u_{\text{sampling}} \), is computed using equation (23).

Here, \( u_{\text{sampling}} \) determines the standard uncertainty contribution originating from the sampling strategy and repeatability of CMM:

\[
 u_{\text{sampling}} = \frac{1}{\sqrt{n_t}} \sqrt{\frac{1}{n_3} \sum_{j=1}^{n_s} \sigma_p^2},
\]

where \( \sigma_p^2 = \frac{1}{n_t} \sum_{j=1}^{n_s} \left( \mu_p - \mu_o \right)^2 \), \( \mu_p = 1/n_3 \sum_{j=1}^{n_s} y_{ij}^p \), and \( \mu_o = 1/n_3 \sum_{j=1}^{n_s} y_{ij}^o \). In which \( \sigma_p^2 \) represents the standard deviation of the \( j \)th evaluation algorithm; i.e., it can be defined as the dispersion caused by varying sampling strategy at the corresponding evaluation algorithm; \( y_{ij}^p \) represents the measured value at the \( j \)th evaluation algorithm, and \( k \)th repetition.

The uncertainty contribution from computation error, \( u_{\text{evaluation algorithm}_i} \), is calculated as the standard deviation of the averages from each evaluation algorithm using the following equation:

\[
 u_{\text{evaluation algorithm}} = \frac{1}{\sqrt{n_s}} \sqrt{\frac{1}{n_3} \sum_{j=1}^{n_s} \left( \mu_p - \mu_o \right)^2} - \frac{u_{\text{sampling}}}{n_t},
\]

where \( \mu_o = 1/n_3 \sum_{j=1}^{n_s} \frac{n_t}{n_s} \sum_{j=1}^{n_s} \mu_R^j \) or \( 1/n_3 \sum_{j=1}^{n_s} \mu_p^j \) and \( \mu_p \) represents the overall mean, i.e., an average of all measured values. In order to avoid overestimation of the combined uncertainty, \( u_{\text{sampling}}/n_t \) is subtracted as in equation (24).

The probing system is one of the crucial components which affects the overall performance of the whole CMM system [104, 105]. The multiple orientation technique is used to compute the uncertainty caused by geometrical errors of the CMM and the probing system [106]. Therefore, the cylindrical workpiece is measured in various positions and orientations using different probe configurations. As shown in Figure 7, three probes of different configurations are selected to quantify measurement uncertainty. These probes exhibit different overall length, sphere diameter, and material. For instance, Probe 1 (overall length–50 mm, sphere diameter–3 mm, and material–rubby), Probe 2 (overall length–75 mm, sphere diameter–5 mm, and material–silicon nitride), and Probe 3 (overall length–33 mm, sphere diameter–3 mm, and material–rubby).

The probing system is specified in terms of the rotation of the probe tip in vertical and horizontal directions as shown in Figure 8. The rotation of the probe about the z-axis characterizes angles in the horizontal plane, whereas rotation about the y-axis assigns angles in the vertical plane [107], as illustrated in Figure 8. For instance, A90° B180° in Figure 8(c) points at the probe, which is oriented at 90° in the vertical plane and 180° in the horizontal plane.

The measurement uncertainty, in this case, has been computed using equations (23) and (24). This measurement uncertainty is the representative of CMM geometric error.
probing error, workpiece positioning error, and fixture error. The assessment of the influence of workpiece surface characteristics on measurement uncertainty is a cumbersome task. This is because the magnitude and the topology of the true workpiece form cannot be understood perfectly. Nevertheless, workpiece shape or surface characteristics can be estimated from the production strategy [4]. As a result, the measurement uncertainty owing to surface conditions is estimated by measuring several workpieces machined under different conditions (like spindle speed and feed rate).
six cylindrical parts are machined under various conditions to quantify the measurement error caused as a result of the workpiece surface. Hence, the uncertainty associated with workpiece characteristics can be computed using equations (25)–(27) [108]:

\[ u_{\text{workpiece}} = \frac{\sigma_p}{\sqrt{n_1}}, \quad (25) \]

\[ \sigma_p = \sqrt{\frac{1}{(n_1 - 1)} \sum_{i=1}^{n_1} (\mu_k - \mu_p)^2}, \quad (26) \]

\[ u_R^i = \frac{1}{n_2} \sum_{k=1}^{n_2} y_k^i, \]

\[ \mu_p = \frac{1}{n_1} \sum_{i=1}^{n_1} \mu_p^i, \quad (27) \]

where \( \sigma_p \) represents the standard deviation, i.e., it can be defined as the dispersion caused by variable surface characteristics; \( y_k^i \) represents the measured value at the \( i^{th} \) machining condition and \( k^{th} \) repetition; \( u_{\text{workpiece}} \) represents the estimated standard deviation and type-A measurement uncertainty introduced as a result of surface properties (material, finish, dust, and lubricant).

The operator should also be considered as one of the crucial factors because it does have a pronounced impact on the measurement uncertainty [24]. The operator affects the measurement results owing to its involvement in part preparation as well as fixing, feature definition, and construction of the coordinate system. Certainly, the measurement accuracy differs due to varied abilities of operators and inadequate implementation of procedures [47, 100]. In this work, the uncertainty due to the operator has been assumed to be included in type-A uncertainties computed for sampling strategy, evaluation method, probing system, and workpiece.

Probe qualification is carried out to gauge the effective radius of the stylus tip, which in turn used to compute the probe tip’s center location with respect to the machine’s coordinate system. It can be attributed to the fact that the probe comes into contact with the surface of the component by its circumference. Before the probe qualification, the reference sphere (calibration sphere) has to be calibrated for its size and position on the machine table. To compute measurement uncertainty associated with the probe qualification procedure, the reference sphere is measured repeatedly and then its standard deviation is computed. Subsequently, the standard deviation in the measurement of a probe’s effective radius is estimated. The formulation in equation (28) is used to determine the measurement uncertainty associated with the probe qualification procedure:

\[ u_{\text{probe qualification}} = \sqrt{u_{\text{reference sphere}}^2 + u_{\text{effective radius}}^2}, \quad (28) \]

where \( u_{\text{reference sphere}} \) represents the estimated standard deviation and type-A measurement uncertainty associated with size measurement of the reference sphere and \( u_{\text{effective radius}} \) represents the standard deviation associated with the measurement of the effective radius of the probe.

Temperature and dust are critical among all the environmental factors, and their effects are often underestimated. For the superior performance of the CMM, the ambient temperature should lie within a specified temperature range and the maximum temperature must not be exceeded. The standard reference temperature of measurement for the current machine is 20°C. It means the geometric errors can be considered constant, at the temperature of 20°C. However, when the temperature fluctuates from 20°C, the errors may be introduced as a result of thermal deformations of the CMM structure. The uncertainty due to ambient temperature variation from the standard temperature of 20°C affects the coefficient of linear expansion(s) of material CMM scales as well as the workpiece. The uncertainty contribution from dust and other dirt particles is presumed to be insignificant. It is because any dust or dirt particles that could intervene with the measuring outcome are cleaned. During experimentation, the laboratory temperature is controlled within 20±0.5°C. The maximum deviation from the standard temperature does not exceed ±0.5°C at any instant in the laboratory. The uncertainty component \( u_{\text{temperature}} \) from temperature, can be computed by assuming that the underlying random deviations are distributed in a rectangular fashion (see equation (29)). The temperature effect is considered through the difference between the coefficient of thermal expansion (CTE) of the CMM scales and the workpiece [27]:

\[ u_{\text{temperature}} = \frac{a}{\sqrt{3}}, \quad (29) \]

where \( a = \text{CTE} \times D \times \Delta T; D = \text{feature size}; \Delta T = \text{temperature variation}; \text{CTE} = \left| \alpha_{\text{CMM}} - \alpha_{\text{workpiece}} \right| ; \alpha_{\text{CMM}} = \text{coefficient of thermal expansion of CMM scales}; \alpha_{\text{workpiece}} = \text{coefficient of thermal expansion of the workpiece}. \)

The overall uncertainty evaluation must address all uncertainty sources in the measurement process. The combined uncertainty (\( u_c \)) combines the respective uncertainty components using the square root of the sum of squares of the individual evaluated uncertainties, as in the following equation:

\[ u_c = \sqrt{u_{\text{sampling}}^2 + u_{\text{evaluation algorithm}}^2 + u_{\text{probe}}^2 + u_{\text{orientation}}^2 + u_{\text{workpiece}}^2 + u_{\text{probe qualification}}^2 + u_{\text{temperature}}^2}, \quad (30) \]
For practical applications, it is important to specify an interval about the measurement result within which the value of the measurand is expected to fall with some confidence. This gives rise to expanded uncertainty, \( U \), which can be computed by multiplying the value of \( u_c \) with a coverage factor \( k \) as in equation (31). A specific value of the coverage factor offers a level of confidence for the expanded uncertainty. For example, a coverage factor of 3 is used to represent a 99.73% confidence interval for the expanded uncertainty. Most often, the overall uncertainty is scaled up utilizing the coverage factor \( k = 2 \), which gives a confidence level of 95.45% (assuming that the combined standard uncertainty is normally distributed):

\[
\text{expanded uncertainty, } U = k \cdot u_c. \tag{31}
\]

5. Results and Discussion

5.1. Predictive Models. This section covers the illustration of the relative performance of different ML algorithms for estimating cylindricity and measurement time. The resulting performance indices and their corresponding parameter setting are presented in Table 2. The comparison of the performance designators indicates that the Gaussian SVR, with the highest prediction accuracy (RMSE ~ 0.000885, MAE ~ 0.000508, and \( R^2 \sim 0.74 \)), is the best regression method for predicting cylindricity. Also, the decision regression trees produce the most accurate predictive model for measurement time. It attains the values of 0.01425, 0.00764, and 1 for RMSE, MAE, and \( R^2 \), respectively. Despite its agility, simplicity, and easiness, the LR model has failed to apprehend the entire complexity of the dataset, especially the nonlinearities that exist in several of the input variables. Indeed, the highly constrained structure of LR models implies that these models commonly have inferior prediction accuracy. As a consequence, more flexible and sophisticated regression models, such as SVR, decision trees, and the ensemble of decision trees, are encouraged. Furthermore, the performance of ensembles of trees (bootstrap aggregating or bagging) is analogous to the performance of SVR in the prognosis of cylindricity. Similarly, the effectiveness of ensembles of trees (least-squares boosting or LSBoost) is close to the performance of decision trees in projecting the measurement time.

The comparison outcomes suggest that the predictive models trained by the SVR (Gaussian) and decision trees (fine) are capable of predicting cylindricity and measurement time, respectively, in real time with the given measurement results. These methods predict cylindricity and measurement time using input parameters such as point distribution strategy, step width, evaluation algorithm, and the number of paths. The refined, practical, and computational exhaustive regression techniques, such as SVR and ensemble of trees, explain only 74% and 72%, respectively, of the variability experienced in the present dataset for cylindricity. This disparity between actual and predicted values can be interpreted through the following reasons. Besides four predictors being used in the investigation, there are apparently several other predictors related to the evaluation of cylindricity. The other predictors may consist of environmental conditions, workpiece properties, probing method, part fixturing, etc. This is why the incorporation of measurement uncertainty is emphasized, particularly in this work. Furthermore, adding more input parameters to the model would increase the size of the dataset and will presumably be useful for cylindricity assessment. Furthermore, the inadequate measurement of cylindricity, owing to manual intervention or limited operator abilities, can also be accountable to the inconsistency in the predicted and true values. Henceforth, the performance of the regression models can be further improved by including more predictors and observations to the dataset, as well as enhancing the quality of data by minimizing the measurement uncertainty.

Regarding the general characteristics of the different regression models, LR can be considered as unreliable due to its low prediction accuracy but a robust technique, whose primary benefit lies in its interpretability and easiness to get trained. It also has high computational speed, in contrast to other advanced techniques. Likewise, the decision trees are fast, low on memory usage, and straightforward and uncomplicated to explain. Moreover, the presence of hyperparameters (usually only the minimum leaf size is needed) is unsubstantial in the case of decision trees. A fine tree with numerous tiny leaves is mostly very precise on the training data. However, a very leafy tree is likely to overfit, and its validation accuracy is usually very low than its training accuracy [109]. Among the major limitations of the decision trees are their complexity. Although they are simple to interpret, constructing a large tree with many branches is complicated and time consuming. The intricacies and higher computational requirements of large trees restrict their extensive application as a regression tool. However, in this study, the decision tree is small as well as

---

### Table 2: Relative performance of various ML algorithms.

| ML algorithms | Best parameters | Cylindricity (mm) | Measurement time (minute) |
|---------------|-----------------|------------------|--------------------------|
|               | RMSE | MAE | \( R^2 \) | RMSE | MAE | \( R^2 \) |
| Linear regression | 0.000962 | 0.000563 | 0.69 | 0.40907 | 0.34142 | 0.89 |
| Decision trees | Minimum leaf size: 4 0.000914 | 0.000553 | 0.72 | 0.01425 | 0.00764 | 1.00 |
| SVR | Kernel function: Gaussian Box constraint: 0.001253 | 0.000885 | 0.000508 | 0.74 | 0.04716 | 0.03928 | 1.00 |
| Ensemble | Method: bagged Minimum leaf size: 1 0.000900 | 0.00052 | 0.72 | 0.01614 | 0.00986 | 1.00 |
|               | | | | | | |
uncomplicated resulting in an explicable and accurate regression model for prediction of measurement time. The SVR belongs to a distinguished family of advanced regression models that provide a significant enhancement over the conventional statistical techniques. However, this higher accuracy is attained on account of higher computational requirements. The nonlinear SVMs are challenging and hard to interpret, but they provide higher predictive accuracy. Additionally, the finer performance of ensemble algorithms demonstrates that these techniques can be examined further.

According to the comparative analysis of different ML approaches, the predictive performance of SVR and fine decision tree can be reported as best for cylindricity and measurement time, respectively. The scatter plot of the actual and predicted values represents one of the leading forms of data visualization. This plot is significant to realize the robustness of the regression model in shaping predictions for various responses. It helps to visually assess the goodness of fit for the developed regression model. The lower $R^2$-squared value and highly dispersed data points suggest an undependable goodness of fit for the model. Certainly, the accuracy of prediction is satisfactory, if the predicted and true responses in the plot are equivalent to the diagonal line. It means that the closer the observations are to the diagonal line, the better is the model. In Figure 9(a), most of the points are lying at close quarters to the diagonal line although there are also some observations with a slight error. It indicates that the developed SVR (Gaussian) model is accurate, and there is a good correlation between the model’s prediction and its true values. Similarly, the residual plot is also employed to assess model performance. It is a scatter plot of residuals on the $y$-axis and actual values on the $x$-axis. This plot is useful to identify nonlinearity, unequal error variances, and outliers in the data. Generally, a good model includes residuals distributed symmetrically and randomly around zero, likely to clutter towards the center of the plot. Moreover, the nonclear pattern or trend in the residuals demonstrates that the regression model does not need further improvement. The plot in Figure 9(b) illustrates that the variation around the zero lines is symmetric and arbitrary indicating that the model is adequate. The foggy and uncertain pattern further emphasizes that the acquired SVR model is ideal to predict cylindricity.

The plot in Figure 10(a) confirms the appropriateness of the fine decision tree model in the prognosis of the measurement time. It represents a perfect regression model for the measurement time, where the predicted response is roughly equal to the true response, and all the points are lying on the diagonal line. The errors are nearly negligible, and the predictions are scattered closer to the diagonal line. Likewise, in the residual plot (Figure 10(b)), the residuals are scattered uniformly and haphazardly around the zero horizontal lines. Moreover, the indistinct and ill-defined pattern in the residuals establishes that the regression model based on fine decision trees is the finest to assess the measurement time.

The experimental and predictive values are acquired for twenty-five randomly generated sequences as presented in Table S1 (in supplementary file) to validate the predictive performance of the Gaussian SVR and fine decision trees. The validation is done to objectively assess the viability and robustness of the models acquired. From the validation plots in Figure 11, it is noticeable that a strong correlation prevails between the predictions of the models and the experimental results. The results of the validation using SVR for cylindricity produce an average change of 2.9 percent between the expected and experimental values. Likewise, the average percentage change of 3.3 percent is observed in the measuring time. The validation experiments confirm that the predictive models developed are robust and adequate for predicting cylindricity and measuring time.
5.2. Multiobjective Optimization. The inspection planning is a multiobjective optimization problem because it consists of two performance characteristics that need to be minimized. The two performance metrics, i.e., cylindricity error and inspection time, are considered to assess the optimal sampling method. The performance criterion, that is, “cylindricity error,” is calculated as the deviation between the cylindricity obtained using the particular sampling strategy and the true value of cylindricity for the given surface. The lower value of error reflects the higher accuracy of the given combination. The true value of 0.0230 mm is obtained for cylindricity using the set of 10,000 points. The inspection time of 31 min 51 sec is recorded when the surface is measured utilizing the greater number of inspection points.

The raw data for GRA execution is obtained by conducting 144 experiments. As the first step in GRA, the linear normalization is carried out, where the original sequence (each sampling strategy) is translated to a comparability sequence using equation (12). The smaller-the-better approach is used because both responses (error and measurement time) are needed to be minimized. All the experimental runs after normalization are denoted as \( s^*_j(l) \) and \( s^*_o(l) \) for comparability and reference sequences, respectively. The difference sequences \( \Delta o_j(l) = |s^*_o(l) - s^*_j(l)| \) are estimated. Next, the GRC is calculated using equations (12) and (13). At the value of \( \xi = 0.9 \), the GRA reported a more consistent outcome with that of FTOPSIS, and hence its value is set as 0.9 in this investigation. It has been assumed that GRA at \( \xi = 0.9 \) represents the actual inspection situation. Kendall’s coefficient of 0.5809 indicates a reasonable agreement between the two optimization methods. The distinguishing coefficient \( \xi = 0.9 \) is substituted in equation (12). Then, the weighted GRG is determined by using equation (14). In this analysis, the error (\( \beta_1 = 0.5 \)) and the measurement time (\( \beta_2 = 0.5 \)) are
assigned equal weights. Table S2 identifies the weighted GRG for each experiment employing the full factorial design. A higher weighted GRG demonstrates that the corresponding responses are comparable to the ideally normalized responses. Experiment 77 is found to have the highest weighted GRG (see Table S2 in the supplementary file); thus, it can be considered the right experimental sequence for optimum results.

The corresponding sequence is point distribution strategy: helix, step width $0.5\text{mm}$, evaluation algorithm $\text{LSQ}$, and number of paths $6$. The following reasons are accountable for achieving better performance with the specified combination:

(i) Helix measurement strategy is not very regular and therefore does not follow systematic or periodic deformations like roundness profiles and generatrix lines strategies. It implies that there is a certain degree of randomness in helix point distribution which is usually beneficial.

(ii) Moreover, the roundness profiles or generatrix lines point distribution exhibit an equally spaced circular or line pattern, around the cylindrical surface. As a result, there is a likelihood that the lobing effect is not recorded in these strategies.

(iii) The combination of the helix measurement strategy, a step width of $0.5\text{mm}$, and a number of paths equal to six leads to uniform and all-inclusive coverage of measurement points over the surface.

(iv) The LSQ in this sequence provides outlier’s robustness, superior computational performance, and the capacity to handle large datasets.

(v) The distribution of the points is certainly an influential variable, and it is therefore advised to prevent even and intense distribution of the points in order to obtain a reasonable and meaningful measurement outcome. There must indeed be a trade-off between the measurement accuracy and the efficiency of the process.

The robustness of the GRA output at $\xi=0.9$ is also analyzed by altering the values of the distinguished coefficient and estimating Kendall’s coefficient. The ranking order of various experiments has identified to be steady, and there is not much variance with fluctuations in $\xi$. In each case, a higher value of $W$ in the range of $0.9940-0.9997$ is observed with the change in $\xi$ from 0.1 to 0.8, and it suggests the accuracy and consistency of the GRA result.

The FTOPSIS approach has also been used for multiobjective optimization to ascertain the results estimated from the GRA approach. The specific intervals (as shown in Table 3) for different linguistic variables are described depending on the accrued data (error and time taken). These linguistic variables, in effect, demonstrate the performance of the sampling strategies and enable to decide the sample size and the point distribution algorithm.

Table 3: Linguistic variables and corresponding fuzzy numbers.

| Variable       | Linguistic values | Range           | Fuzzy numbers |
|----------------|-------------------|-----------------|---------------|
| Error          |                    |                 |               |
| Very low       |                   | $\leq 0.0008$   | (0.0005, 0.0006, 0.0007, 0.0008) |
| Low            |                   | 0.0008–0.002    | (0.0008, 0.00012, 0.00016, 0.002) |
| Middle         |                   | 0.002–0.008     | (0.002, 0.004, 0.006, 0.008) |
| High           |                   | 0.008–0.011     | (0.008, 0.009, 0.01, 0.011) |
| Very high      |                   | $> 0.011$       | (0.011, 0.051, 0.091, 0.13) |
| Measurement time |                  |                 |               |
| Very low       |                   | $\leq 0.4$      | (0.1, 0.2, 0.3, 0.4) |
| Low            |                   | 0.41–0.70       | (0.41, 0.5, 0.6, 0.70) |
| Middle         |                   | 0.71–1          | (0.71, 0.8, 0.9, 1) |
| High           |                   | 1.1–3.5         | (1.01, 1.85, 2.7, 3.55) |
| Very high      |                   | $> 3.55$        | (3.551, 4.55, 5.55, 6.55) |

Figure 12: Performance at optimized and randomly selected parameters.
The crisp values (data acquired) are transformed into trapezoidal fuzzy numbers in the first step of FTOPSIS using the fuzzy numbers shown in Table 3. Next, using equations (15)–(18), the normalized and weighted standard decision matrices are generated from the matrix of the fuzzy numbers. It is accompanied by the FPIS and FNIS computations using equation (19). Then, the distance of each combination from FPIS and FNIS is computed using equation (20), and finally the CC is estimated using equation (21). Notice that the responses are given equal priority (that is, weights of 0.5 each). Table S3 (in the supplementary file) displays the values of CC and respective ranks for each experiment. A higher CC implies that the responses are nearer to the optimally normalized responses. Experiment 77 is noticed to provide the highest CC (see Table S3 in the supplementary file); thereby, it can be specified as the leading experimental sequence or the ideal sampling strategy. The corresponding sequence is point distribution strategy = helix, step width = 0.5 mm, evaluation algorithm = LSQ, and number of paths = 6, which is the same as GRA has provided.

After obtaining the appropriate sampling strategy, the validation is carried out by repeating the experiments at optimized and randomly chosen parameters. As can be seen in Figure 12, the perfect balance between accuracy and measuring efficiency can only be achieved with optimized scanning parameters. Although it is possible to obtain the lowest time or the highest accuracy at random parameters, they cannot provide consistency between efficiency and precision.

### 5.3. Uncertainty Assessment

The different components of measurement uncertainty sources associated with the assessment of cylindricity can be realized in the uncertainty budget as shown in Table 4. It reports quantified measurement uncertainty values and their various components, as well as combined and expanded uncertainties. The components can be described as the intrinsic elements which constitute the identified measuring uncertainty. For instance, components such as geometric error, workpiece positioning, probe orientation, fixture, and operator glitch constitute the gross measuring uncertainty

#### Table 4: Uncertainty budget in relation to cylindricity evaluation using CMM.

| Factors               | Components                          | Measurement uncertainty |
|-----------------------|-------------------------------------|-------------------------|
|                       |                                     | Type A                  |
|                       |                                     | $u_{sampling}$  | $u_{evaluation}$ algorithm | $u_{probe}$ | $u_{orientation}$ | $u_{workpiece}$ | $u_{probe qualification}$ | $u_{temperature}$ |
| Machine axes/geometric error |                                     |                         |                         |             |                 |                 |                      |                 |
| Scales temperature    |                                     |                         |                         |             |                 |                 |                      |                 |
| Probe—ball diameter   |                                     |                         |                         |             |                 |                 |                      |                 |
| Styli length          |                                     |                         |                         |             |                 |                 |                      |                 |
| Probe type (ball material) |                                 |                         |                         |             |                 |                 |                      |                 |
| Probe orientation     |                                     |                         |                         |             |                 |                 |                      |                 |
| Probe qualification   |                                     |                         |                         |             |                 |                 |                      |                 |
| Number of points      |                                     |                         |                         |             |                 |                 |                      |                 |
| Point distribution    |                                     |                         |                         |             |                 |                 |                      |                 |
| Evaluation algorithm  |                                     |                         |                         |             |                 |                 |                      |                 |
| Surface quality       |                                     |                         |                         |             |                 |                 |                      |                 |
| Orientation           |                                     |                         |                         |             |                 |                 |                      |                 |
| Positioning           |                                     |                         |                         |             |                 |                 |                      |                 |
| Fixture               |                                     |                         |                         |             |                 |                 |                      |                 |
| Form                  |                                     |                         |                         |             |                 |                 |                      |                 |
| Size                  |                                     |                         |                         |             |                 |                 |                      |                 |
| Temperature           |                                     |                         |                         |             |                 |                 |                      |                 |
| Alignment             |                                     |                         |                         |             |                 |                 |                      |                 |
| Clamping              |                                     |                         |                         |             |                 |                 |                      |                 |
| Part preparation      |                                     |                         |                         |             |                 |                 |                      |                 |
| Feature definition    |                                     |                         |                         |             |                 |                 |                      |                 |
| Temperature           |                                     |                         |                         |             |                 |                 |                      |                 |
| Dust                  |                                     |                         |                         |             |                 |                 |                      |                 |
| Evaluated uncertainty |                                     | 0.00028                 | 0.000098                | 0.000497    | 0.000433       | 0.000400        | 0.000790              | 0.000080          |
| Combined uncertainty  |                                     | 0.001147                |                         |             |                 |                 |                      |                 |
| Coverage factor       |                                     |                         |                         |             |                 |                 |                      | 2                 |
| Expanded uncertainty  |                                     |                         |                         |             |                 |                 |                      | 0.02294           |
estimated for part orientation. The measurement uncertainties due to sampling strategy and evaluation algorithm calculated using equations (23) and (24) are 0.000284 and 0.000098 mm, respectively. Similarly, the measurement uncertainties due to probe and its orientation are found to be 0.000497 and 0.000433 mm, respectively. Using equations (25)–(27), the uncertainty associated with workpiece characteristics is computed as 0.00040 mm. The probe qualification procedure and the temperature also contributed to the overall uncertainty, and their values are acquired using equations (28) and (29), respectively. Eventually, the combined and expanded uncertainties computed using equations (30) and (31) are 0.001147 and 0.002294 mm. The complete statement of the true value of the cylindricity can be provided after the uncertainty assessment as 0.0223 ± 0.002294 mm.

6. Conclusion

A dependable sampling plan is necessary if an accurate inspection is to be carried out in minimal time, effort, and resources. The accomplishment of the effective sampling strategy and the usage of relevant evaluation algorithms are fundamental problems in the CMM inspection. Multiple solutions have been developed to define the sampling strategy, but literally, research has been undertaken to reliably predict cylindricity and inspection time using ML techniques. In addition, comprehensive research using GRA and FTOPSIS to optimize the inspection strategy is deficient in previous works. This research, therefore, emphasizes on the deployment of predictive models and identification of an optimal sampling plan which alleviate the cost of the inspection, as well as ensures measuring reliability. The best prediction performance has been observed using Gaussian SVR and free decision trees for cylindricity and measurement time, respectively. The confirmation experiments acknowledge the robustness and adequacy of the predictive models developed to predict cylindricity and measurement time. The validation outcomes using SVR for cylindricity yields an average change of 2.9 percent between the expected and experimental values. Similarly, an average change of 3.3 percent is recorded for the measuring time. The allocation of points is unquestionably a prominent factor, and thus it is prudent to inhibit an even and rigorous distribution of the points in order to establish a rational and substantive measuring result. In this research, the multiobjective optimization approaches successfully provide an inspection plan that has achieved a compromise between measurement accuracy and efficiency. The acquired combination is point distribution strategy = helix, step width = 0.5 mm, evaluation algorithm = LSQ, and number of paths = 6. Both GRA and FTOPSIS provide the same combination, which confirms that the acquired sampling strategy is the most suitable to evaluate the cylindricity. Furthermore, the verification analysis also reports that these scanning parameters can pursue the right balance between precision and time. Certainly, the assessment of measurement uncertainty has been an important issue in coordinate metrology. The measurement results devoid of uncertainty statements carry little or no practical significance to industries. Literally, it is daunting to infer analogous results on CMM given the presence of multiple known and unknown influencing factors. Nevertheless, to ensure consistency among measurement results, the idea of measurement uncertainty has been incorporated. Finally, the true value of cylindricity has been estimated to be 0.0223 ± 0.00229 mm.

It can also be seen that while metrology research has continued to evolve consistently, there is still plenty of potential for improvement. For example, it will be very useful to develop practical, reliable, easy to achieve, and standard practices for estimating measurement uncertainty of CMMs, as well as to propose automated sampling strategies that can improve versatility, speed, accuracy, etc. The authors in this work are striving to execute more ML strategies in future work, such as ANN. The other multiobjective optimization techniques for the computational performance analysis will be considered. The implemented methodology will also be extended for other form tolerances to expand its applicability and universality.

Data Availability

All of the data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group number RG-1440-026.

Supplementary Materials

Table S1: validation experiment sequence. Table S2: grey relational grade (GRG) for all experimental runs. Table S3: closeness coefficient (CC) for all experimental runs. (Supplementary Materials)

References

[1] D. F. Elkott, H. A. Elmaraghy, and W. H. Elmaraghy, “Automatic sampling for CMM inspection planning of free-form surfaces,” International Journal of Production Research, vol. 40, no. 11, pp. 2653–2676, 2002.
[2] J. M. Baldwin, K. D. Summerhays, D. A. Campbell, and R. P. Henke, “Application of simulation software to coordinate measurement uncertainty evaluations,” NCSLI Measure, vol. 2, no. 4, pp. 40–52, 2007.
[3] A. Weckenmann, M. Knauer, and T. Killmaier, “Uncertainty of coordinate measurements on sheet-metal parts in the automotive industry,” Journal of Materials Processing Technology, vol. 115, no. 1, pp. 9–13, 2001.
[4] A. Weckenmann, M. Knauer, and H. Kunzmann, “The influence of measurement strategy on the uncertainty of CMM-Measurements,” CIRP Annals, vol. 47, no. 1, pp. 451–454, 1998.
[5] G. Moroni and S. Petro, "Coordinate measuring machine measurement planning," "Coordinate measuring machine measurement planning," in Geometric Tolerances, B. M. Colosimo and N. Senin, Eds., Springer, London, UK, 2011.

[6] W. Choi, T. R. Kurfess, and J. Cagan, "Sampling uncertainty in coordinate measurement data analysis," Precision Engineering, vol. 27, no. 3, pp. 153–163, 1998.

[7] G. Lee, J. Mou, and Y. Shen, "Sampling strategy design for dimensional measurement of geometric features using coordinate measuring machine," International Journal of Machine Tools and Manufacture, vol. 37, no. 7, pp. 917–934, 1997.

[8] F. M. M. Chan, T. G. King, and K. J. Stout, "The influence of sampling strategy on a circular feature in coordinate measurements," Measurement, vol. 19, no. 2, pp. 73–81, 1996.

[9] B. M. Colosimo, E. G. Moya, G. Moroni, and S. Petro, "Statistical sampling strategies for geometric tolerance inspection by CMM," Economic Quality Control, vol. 23, no. 1, pp. 109–121, 2008.

[10] T. Weber, S. Motavalli, B. Fallahi, and S. H. Cheraghi, "A unified approach to form error evaluation based on the error evaluation approach to form error evaluation," Precision Engineering, vol. 26, no. 3, pp. 269–278, 2002.

[11] A. Barari and S. Mordo, "Effect of sampling strategy on uncertainty and precision of flatness inspection studied by dynamic minimum deviation zone evaluation of sampling strategy on uncertainty and precision of flatness inspection studied by dynamic minimum deviation zone evaluation," International Journal of Metrology and Quality Engineering, vol. 4, no. 1, pp. 3–8, 2013.

[12] G. Moroni and S. Petro, "Geometric tolerance evaluation: a discussion on minimum zone fitting algorithms," Precision Engineering, vol. 32, no. 3, pp. 232–237, 2008.

[13] Q. Jiang, H.-Y. Feng, D. OuYang, and M. T. Desta, "A roundness evaluation algorithm with reduced fitting uncertainty of CMM measurement data," Journal of Manufacturing Systems, vol. 25, no. 3, pp. 184–195, 2006.

[14] Mathworks, Regression Learner App, https://in.mathworks.com/help/stats/regression-learner-app.html, 2020.

[15] A. Meo, L. Profumo, A. Rossi, and M. Lanzetta, "Optimum dataset size and search space for minimum zone roundness evaluation by genetic algorithm," Measurement Science Review, vol. 13, no. 3, pp. 100–107, 2013.

[16] C.-H. Menq, H.-T. Yau, and G.-Y. Lai, "Automated precision measurement of surface profile in CAD-directed inspection," IEEE Transactions on Robotics and Automation, vol. 8, no. 2, pp. 268–278, 1992.

[17] B. C. Jiang and S.-D. Chiu, "Form tolerance-based measurement points determination with CMM," Journal of Intelligent Manufacturing, vol. 13, no. 2, pp. 101–108, 2002.

[18] R. Edgeworth and R. G. Wilhelm, "Adaptive sampling for coordinate metrology," Precision Engineering, vol. 23, no. 3, pp. 144–154, 1999.

[19] H. Lee, M.-W. Cho, G.-S. Yoon, and J.-H. Choi, "A computer-aided inspection planning system for on-machine measurement-part I: global inspection planning," KSME International Journal, vol. 18, no. 8, pp. 1349–1357, 2004.

[20] M.-W. Cho, H. Lee, G.-S. Yoon, and J.-H. Choi, "A computer-aided inspection planning system for on-machine measurement - Part II: local inspection planning," KSME International Journal, vol. 18, no. 8, pp. 1358–1367, 2004.

[21] M.-W. Cho, H. Lee, G.-S. Yoon, and J. Choi, "A feature-based inspection planning system for coordinate measuring machines," The International Journal of Advanced Manufacturing Technology, vol. 26, no. 9–10, pp. 1078–1087, 2005.

[22] G. Moroni and S. Petro, "State-of-the-art research and future perspectives in CMM metrology," Proceedings of the CIRP Conference on Intelligent Maintenance Systems, 2012.

[23] M. A. Badar, S. Raman, and P. S. Pulat, "Intelligent search-based selection of sample points for straightforward and flatness estimation," ASME Journal of Manufacturing Science and Engineering, vol. 125, no. 2, pp. 263–271, 2003.

[24] C.-X. J. Feng, A. L. Saal, J. G. Salisbury, A. R. Ness, and G. C. S. Lin, "Design and analysis of experiments in CMM measurement uncertainty study," Precision Engineering, vol. 31, no. 2, pp. 94–101, 2007.

[25] P. Pedone and D. Romano, "Designing small samples for form error estimation with coordinate measuring machines," Precision Engineering, vol. 35, no. 2, pp. 262–270, 2011.

[26] H. J. Pakh, M. Y. Jung, S. W. Hwang, Y. H. Kim, Y. S. Hong, and S. G. Kim, "Integrated precision inspection system for manufacturing of moulds having CAD defined features," The International Journal of Advanced Manufacturing Technology, vol. 10, no. 3, pp. 198–207, 1995.

[27] I. Vrba, R. Palenčar, M. Hadžištević, B. Štrbac, and J. Hodolič, "The influence of the sampling strategy and the validation method on the cylindricity error on a coordinate measurement machine," Journal of Production Engineering, vol. 16, no. 2, pp. 53–56, 2013.

[28] P. Guan, T. Yu, and W. Wang, "Research on flatness error measurement of revolving body end-face," Applied Mechanics and Materials, vol. 44, pp. 4002–4006, 2011.

[29] P. Tufekci, "Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods," International Journal of Electrical Power & Energy Systems, vol. 60, pp. 126–140, 2014.

[30] H. A. Guvenir and I. Üysal, "Regression on feature projections," Knowledge-Based Systems, vol. 13, no. 4, pp. 207–214, 2000.

[31] Z. Li, Z. Zhang, J. Shi, and D. Wu, "Prediction of surface roughness in extrusion-based additive manufacturing with machine learning," Robotics and Computer-Integrated Manufacturing, vol. 57, pp. 488–495, 2019.

[32] G. Tapia, A. H. Elwany, and H. Sang, "Prediction of porosity in metal-based additive manufacturing using spatial Gaussian process models," Additive Manufacturing, vol. 12, pp. 282–290, 2016.

[33] J. Batorynska, O. S. Semeniuta, and K. Martiisen, "Optimization of process parameters for powder bed fusion additive manufacturing by combination of machine learning and finite element method: a conceptual framework," Procedia CIRP, vol. 67, pp. 227–232, 2018.

[34] D. Wu, C. Jennings, J. Terpenny, R. X. Gao, and S. Kumara, "A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests," Journal of Manufacturing Science & Engineering, vol. 139, no. 7, 9 pages, Article ID 071018, 2017.

[35] A. Gouarir, G. Martínez-Arellano, G. Terrazas, P. Benardos, and S. Ratchev, "In-process tool wear prediction system based on machine learning techniques and force analysis," Procedia CIRP, vol. 77, pp. 501–504, 2018.

[36] P. Mohammadi and Z. J. Wang, "Machine learning for quality prediction in abrasion-resistant material manufacturing process," in Proceedings of the 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Vancouver, BC, Canada, May 2016.
[71] C.-T. Chen, “Extensions of the TOPSIS for group decision-making under fuzzy environment,” Fuzzy Sets and Systems, vol. 114, no. 1, pp. 1–9, 2000.

[72] L. Shen, L. Olfat, K. Govindan, R. Khodaverdi, and A. Diabat, “A fuzzy multi criteria approach for evaluating green supplier’s performance in green supply chain with linguistic preferences,” Resources, Conservation and Recycling, vol. 74, pp. 170–179, 2013.

[73] A. Awasthi, S. S. Chauhan, and S. K. Goyal, "A multi-criteria decision making approach for location planning for urban distribution centers under uncertainty," Mathematical and Computer Modelling, vol. 53, no. 1-2, pp. 98–109, 2011.

[74] S. K. Pattnaik, M. Priyadarshini, K. D. Mahapatra, D. Mishra, and S. Panda, "Multi objective optimization of EDM process parameters using FUZZY TOPSIS method," in Proceedings of the 2nd International Conference on Innovations in Information Embedded and Communications Systems (ICIIECS), Coimbatore, India, March 2015.

[75] N. Weihua and Y. Zhenqiang, "Cylindricity modeling and tolerance analysis for cylindrical components," The International Journal of Advanced Manufacturing Technology, vol. 64, no. 5–8, pp. 867–874, 2013.

[76] L. Xianqing, S. Hongwei, X. Yujun, L. Jishun, Z. Jia, and D. Mingde, "Method for cylindricity error evaluation using geometry optimization searching algorithm," Measurement, vol. 44, pp. 1556–1563, 2011.

[77] Standard ISO 12180-1: 2011, Geometrical Product Specifications (GPS)—Cylindricity—Part 1: Vocabulary and Parameters of Cylindrical Form, International Organization for Standardization, Geneva, Switzerland, 2011.

[78] M. Dovica and J. Buša, “Method comparison of the evaluation of the cylindricity deviation measured by different measurement strategies,” Acta Mechanica Slovaca, vol. 17, no. 3, pp. 94–97, 2013.

[79] Standard ISO 12180-2: 2011, Geometrical Product Specifications (GPS)—Cylindricity—Part 2: Specification Operators, International Organization for Standardization, Geneva, Switzerland, 2011.

[80] B. Gapinski, A. Kolodziej, M. Pawlowski, L. Marciniak, M. Grzelka, and M. Lukaszyk, “Influence of measurement strategy on the value of cylindricity,” in Proceedings of the XX IMEKO World Congress Metrology for Green Growth, Busan, Republic of Korea, September 2012.

[81] O. Devillers and F. Preparata, "Evaluating the cylindricity of a nominally cylindrical point set," in Proceedings of the 11th ACM-SIAM Symposium on Discrete Algorithms, San Francisco, CA, USA, January 2000.

[82] M. Yu, Y. Zhang, Y. Li, and D. Zhang, “Adaptive sampling method for inspection planning on CMM for free-form surfaces,” The International Journal of Advanced Manufacturing Technology, vol. 67, no. 9–12, pp. 1967–1975, 2013.

[83] A. Rossi and M. Lanzetta, "Optimal blind sampling strategy for minimum zone roundness evaluation by metaheuristics," Precision Engineering, vol. 37, no. 2, pp. 241–247, 2013.

[84] D. D. Diren, S. Bora, I. H. Selvi, and T. Hatipoglu, Root Cause Detection with an Ensemble Machine Learning Approach in the Multivariate Manufacturing Process, Industrial Engineering in the Big Data Era, 2016, F. Calisir, E. Cevikcan, and H. C. Akdag, Eds., Springer, Cham, Switzerland, 2019.

[85] E. L. Lehmann and G. Casella, Theory of Point Estimation, Springer-Verlag, New York, NY, USA, 2nd edition, 1998.

[86] M. Kuhn, “Building predictive models in R using the caret package,” Journal of Statistical Software, vol. 28, no. 5, pp. 1–26, 2008.

[87] Data Driven Investor, K-Fold Cross Validation, https://medium.com/datadriveninvestor/k-fold-cross-validation-6b8518070833, 2020.

[88] H. S. Lu, C. K. Chang, N. C. Hwang, and C. T. Chung, "Grey relational analysis coupled with principal component analysis for optimization design of the cutting parameters in high-speed end milling." Journal of Materials Processing Technology, vol. 209, no. 8, pp. 3808–3817, 2009.

[89] W.-S. Lee and Y.-C. Lin, "Evaluating and ranking energy performance of office buildings using Grey relational analysis," Energy, vol. 36, no. 5, pp. 2551–2556, 2011.

[90] L. F. Wu, S. F. Liu, L. G. Yao, and S. L. Yan, "Grey convex relational degree and its application to evaluate regional economic sustainability," Scientia Iranica, vol. 20, no. 1, pp. 44–49, 2013.

[91] Y. Kuo, T. Yang, and G.-W. Huang, "The use of grey relational analysis in solving multiple attribute decision-making problems," Computers & Industrial Engineering, vol. 55, no. 1, pp. 80–93, 2008.

[92] H.-H. Wu, "A comparative study of using grey relational analysis in multiple attribute decision making problems," Quality Engineering, vol. 15, no. 2, pp. 209–217, 2002.

[93] C.-T. Chen, C.-T. Lin, and S.-F. Huang, "A fuzzy approach for supplier evaluation and selection in supply chain management," International Journal of Production Economics, vol. 102, no. 2, pp. 289–301, 2006.

[94] S. Nadaban, S. Dritac, and I. Dritaca, "Fuzzy TOPSIS: a general view," Procedia Computer Science, vol. 91, pp. 823–831, 2016.

[95] M. Savitha and G. Mary, "New methods for ranking of trapezoidal fuzzy numbers," Advances in Fuzzy Mathematics, vol. 12, pp. 1159–1170, 2017.

[96] G.-S. Liang and M.-J. J. Wang, "A fuzzy multi-criteria decision-making method for facility site selection," International Journal of Production Research, vol. 29, no. 11, pp. 2313–2330, 1991.

[97] A. Habibi, A. Sarafrazi, and S. Izadyar, "Delphi technique theoretical framework in Qualitative research," The International Journal of Engineering and Science, vol. 3, pp. 8–13, 2014.

[98] R. Schmidt, K. Lyytinen, M. Keil, and P. Cule, "Identifying software project risks: an international Delphi study," Journal of Management Information Systems, vol. 17, no. 4, pp. 5–36, 2001.

[99] J. A. Sładek, "Chapter: 2-measurement uncertainty and reliability of characterization," in Measurement Uncertainty in Qualitative Research, Springer, Berlin, Heidelberg, Germany, 2016.

[100] P. Saunders, A. Wilson, N. Orchard, N. Tatman, and P. Maropoulos, "An exploration into measurement consistency on coordinate measuring machines," Procedia CIRP, vol. 25, pp. 19–26, 2014.

[101] G1owy Urząd Miar—Central Office of Measures, Guide to the Expression of Uncertainty in Measurement (GUM) ISO 1993/1995 with the Addition to Polish Edition of J. M. Jaworski, Central Office of Measures, Warsaw, Poland, 1999.

[102] A. Elahdi, M. Yahya, and M. Halaj, "Uncertainty and its impact on the quality of measurement," American Journal of Engineering and Applied Sciences, vol. 5, no. 2, pp. 114–118, 2012.
[103] ISO/TS 15530-2:2004, Geometrical Product Specifications (GPS)–Coordinate Measuring Machines (CMM): Technique for Determining the Uncertainty of Measurement–Part 2: Use of Multiple Measurement Strategies, ISO, Geneva, Switzerland, 2004.

[104] A. Wozniak and M. Dobosz, “Metrological feasibilities of CMM touch trigger probes. part I: 3D theoretical model of probe pre-travel,” Measurement, vol. 34, no. 4, pp. 273–286, 2003.

[105] P. A. Cauchick-Miguel and T. G. King, “Factors which influence CMM touch trigger probe performance,” International Journal of Machine Tools and Manufacture, vol. 38, no. 4, pp. 363–374, 1998.

[106] S. Osawa, K. Busch, M. Franke, and H. Schwenke, “Multiple orientation technique for the calibration of cylindrical workpieces on CMMs,” Precision Engineering, vol. 29, no. 1, pp. 56–64, 2005.

[107] S. H. Mian, M. A. Mannan, and A. M. Al-Ahmari, “The influence of surface topology on the quality of the point cloud data acquired with laser line scanning probe,” Sensor Review, vol. 34, no. 3, pp. 255–265, 2014.

[108] S. Castrup, “Comparison of methods for establishing confidence limits and expanded uncertainties,” in Proceedings of the 2010 Measurement Science Conference, pp. 1–23, Pasadena, CA, USA, 2010.

[109] Mathworks, Choose Regression Model Options, https://in.mathworks.com/help/stats/choose-regression-model-options.html, 2020.