# Integrating Spatial Statistics and Digital Processing for Enhance Surface Quality Classification in the Foundry Industry

Ronit Shetty, Ahmad Al Majali & Lee Wells
Western Michigan University, Kalamazoo, Michigan, USA



## ABSTRACT

To ensure high quality castings, the ability to accurately quantify an as-cast surface's characteristics is of vital importance to the foundry industry. In addition, recent advancements in non-contact measurement systems have provided new opportunities to quantify a casting's surface beyond traditional roughness measurements. However, in the realm of non-contact measurement systems, there are numerous methodologies and metrics for evaluating and quantifying a surface. This paper investigates the critical surface features and approaches necessary for effective surface quality classification. More specifically, this study compares the accuracy of spatial statistical, modern digital processing techniques (e.g., convolution neural network) to classify as-cast surfaces. Through this comparison, this paper aims to develop a methodology to provide the most reliable results to industry. The findings will help guide the foundry industry in adopting the most appropriate techniques for surface quality assessment, ultimately enhancing product quality.

**Keywords**: variogram, multi-way principal component analysis, MPCA, surface roughness, vectorized-principal component analysis, VPCA, machine learning

## LITERATURE REVIEW

The foundry industry relies heavily on precise surface quality assessments to ensure high-quality castings. With recent advancements in non-contact measurement systems and machine learning, there are now more advanced techniques for evaluating and quantifying surface features. These techniques allow for a more comprehensive understanding of surface characteristics, especially in complex manufacturing processes like casting, where surface quality is critical for functionality and performance.

This literature review explores a range of methodologies for surface quality classification, highlighting the shift from conventional roughness measurements to more sophisticated digital processing techniques. It examines key concepts such as surface texture, point cloud data, and modern machine learning models like Convolutional Neural Networks (CNNs), alongside statistical approaches for dimensionality reduction, such as Multi-Way Principal Component Analysis (MPCA).

### SURFACE ROUGHNESS

Surface texture refers to the recurring or irregular variations from the intended surface, which together form the three-dimensional topography of the surface. This term encompasses several key elements, including roughness (at nano- and micrometer scales), waviness (at the macroscopic level), lay, and flaws.[1]

Surface roughness, the primary classification focus of this paper, plays a crucial role in manufacturing systems, influencing the functionality and performance of materials and components. Nano- and microroughness refer to surface variations at very short wavelengths, appearing as hills (asperities) and valleys (local minima) with differing amplitudes and spacings. These features, though significantly larger than molecular dimensions, are integral to surface characteristics. In two dimensions, asperities are seen as peaks, while in three-dimensional surface maps, they are identified as summits. Nano- and microroughness also include features resulting from the manufacturing process, such as traverse feed marks and other irregularities within the roughness sampling length.[2]

#### Arithmetic mean height ($S_a$):

The arithmetic mean height is the three-dimensional counterpart to the two-dimensional average roughness ($R_a$). $S_a$, also known as the "real average roughness," represents the mean height of all points measured over the surface area. While the "R" parameters describe roughness based on a linear profile, the "S" parameters pertain to measurements taken across a surface area. $S_a$ quantifies the surface roughness by calculating the average absolute differences between the height of each point on the surface and the mean height.[3] This metric is widely used in assessing surface roughness characteristics and can be mathematically expressed in Eqn. 1:

$$S_a = \frac{1}{A}\iint_A |z(x,y)|dxdy \qquad \textbf{Eqn.1}$$

Where $A$ is the total surface area and $z(x,y)$ represents the height at any given point on the surface.

### POINT CLOUD

Point clouds are extensive collections of 3D data points that represent the surfaces of objects or scenes, capturing their geometric characteristics and surface qualities, including

finish and roughness. They can be organized in regular or irregular grid formats. These data points are usually produced using methods such as LiDAR, photogrammetry, or structured light scanning.[4]

Recent studies have investigated the integration of point cloud data with machine learning techniques, which significantly improves object recognition, semantic segmentation, and scene understanding.[5] A functional morphing control charting system was introduced by Zhou et al.[6] for evaluating the surface quality of high-definition (HD) data. Wang et al.[7] introduced a Hotelling control charting framework for monitoring point cloud data from semiconductor wafers. This approach facilitates the evaluation of data quality and consistency. They developed an EWMA control chart that utilized features extracted from a Q-Q plot to detect deviations from a nominal surface.

In a study conducted by Helle and Lemu,[8] The application of a handheld 3D laser scanner to a 3D-printed part highlights the effectiveness of 3D scanning in reverse engineering (RE) and production control. The research shows that 3D scanning is particularly beneficial for reverse engineering complex shapes, such as those with detailed and irregular surfaces, which are difficult to measure accurately by traditional methods.

To identify surface flaws from a point cloud, Dastoorian et al.[9] employed the adaptive generalized likelihood ratio test to identify surface defects within point cloud data. Point clouds are becoming increasingly important for quality control in additively manufactured (AM) parts, complementing their use in traditional production methods. This technique proves to be useful for quality assessment in both conventional and additive manufacturing processes. A spectral graph theory (SGT) method for evaluating the surface and geometric integrity of AM components was put out by Rao et al. and Tootooni et al.[10, 11]

## FEATURE ENGINEERING

Feature engineering involves transforming raw data into relevant features that enhance the performance of machine learning models. This process includes selecting, modifying, and creating new variables from the existing data to better capture key patterns and relationships. Techniques such as scaling, encoding categorical variables, and generating feature interactions are commonly used. By refining the data, feature engineering enables models to learn more efficiently and improves their accuracy and predictive power. The following two sections provide an explanation of two key feature engineering techniques used in this paper: Multilinear Principal Component Analysis (MPCA) and the variogram method.

### Multilinear Principal Component Analysis (MPCA)

Low-rank decomposition methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are widely used to characterize high-dimensional data.[12] While PCA-based methods offer benefits in multivariate processing, they encounter limitations when used for image-based process monitoring. Traditional PCA overlooks the spatial correlations between pixel intensities in images, treating them as independent variables. To overcome this limitation, extensions such as 2DPCA have been developed. This two-dimensional PCA method is tailored for image analysis, working directly with image matrices to extract features, rather than first converting them into 1D vectors.[13] For tensor objects that include both spatial and temporal information from images or video frames, two PCA variants—T-mode PCA and S-mode PCA—were introduced.[14]

Colosimo and Grasso [15] enhanced the T-mode PCA methodology by incorporating considerations for the spatial relationships between pixels within an image. Another approach for dimensionality reduction in tensor data is Vectorized-PCA (VPCA), which adapts PCA principles to multi-view arrays by converting them into vectors. However, it is crucial to recognize that this unfolding process results in the loss of spatial correlation information across different modes. To address this limitation, Lu et al.[16] introduced Multilinear principal component analysis (MPCA), a method that can be applied directly to tensors without requiring unfolding.

MPCA has proven to be an effective tool for feature extraction from tensors and is widely used in various process monitoring applications. For instance, Shetty et al. (2024) introduced a novel method for classifying surface roughness by combining image and point cloud data. Using MPCA analysis and a random forest classifier, they developed a model that outperforms one based solely on image data. The study demonstrates that the fused data model provides greater precision and a more comprehensive understanding of surface texture.[17]

Multilinear Principal Component Analysis (MPCA) is an advanced dimensionality reduction technique designed for multi-way data, typically represented as tensors. It operates by unfolding tensors along specific modes, each corresponding to a different dimension of the data. Conventional PCA methods are then applied to these unfolded tensors to extract projection matrices for each mode. The primary objective of MPCA is to maximize the scatter of the projected tensor while maintaining the inherent structure and relationships within the data. The projection matrices are iteratively adjusted until they meet a specified convergence criterion.

### Variogram

The variogram roughness method is an advanced digital technique for measuring surface roughness, particularly effective for castings. Unlike traditional methods focused solely on height variations, this approach uses point cloud data to assess the spatial relationships between surface points. It relies on a variogram, a statistical tool that

measures spatial autocorrelation, or the similarity between points based on their separation distance. This allows for a more detailed analysis of roughness, considering both height differences and spatial distribution.[18]

A variogram, widely used in geostatistics, offers a graphical representation of how spatial correlation changes with distance, providing insight into the structure of a surface.[19] Two key parameters derived from the variogram are the sill and the range. The sill represents the point where semivariance reaches a plateau, indicating that beyond this distance, there is no spatial correlation between points. In other words, the sill defines the limit where the surface's texture no longer exhibits autocorrelation. The range is the distance over which surface points remain correlated, marking the spatial extent of the surface pattern. Once the distance between points exceeds the range, they become independent of each other in terms of roughness.

The method begins by comparing surface points to a reference grid to update their z-values, which are then used to calculate roughness. Its ability to distinguish between surfaces with similar height values but different spatial arrangements make it highly effective at detecting subtle variations that traditional methods might miss. The evaluation length, which determines the averaging range for variogram points, plays a crucial role. Shorter lengths reduce roughness values, while longer ones enhance sensitivity, allowing for a flexible and detailed surface analysis.

Unlike conventional 2D parameters such as $S_a$ or $R_a$, the variogram method integrates both height and spatial data, offering a more comprehensive and objective measure of surface roughness. It also effectively identifies irregularities and deviations from expected surface profiles, making it a reliable tool for inspecting cast or machined surfaces.[20]

## FEATURE ENGINEERING

Supervised classification is one of the most common tasks in Intelligent Systems, leading to the development of various techniques grounded in Artificial Intelligence, such as logic-based and perceptron-based methods, as well as Statistical techniques like Bayesian networks and instance-based approaches. The core goal of supervised learning is to build a concise model that captures the relationship between predictor features and class labels. This trained model, or classifier, is then used to predict class labels for new instances, where the values of the predictor features are available, but the corresponding class labels are unknown.[21]

Machine learning models have become indispensable tools across diverse domains, driving advancements in data-driven decision-making. For instance, Shetty et al. (2024) introduced a thorough approach for classifying refractory coatings used on chemically bonded sand according to their thickness, which is essential for monitoring mold and core coatings in foundries. The method combines feature extraction through vectorized principal component analysis (VPCA) with classification modeling using a machine learning algorithm. The study demonstrated the efficacy of feature extraction methods and machine learning algorithms in accurately categorizing coating thicknesses, providing practical solutions for applications in the foundry industry 29. In this context, a concise overview of the machine learning models used in this paper, namely Decision Trees and Support Vector Machines (SVMs), is presented.[22]

### Decision Trees

Decision Trees are a foundational concept in machine learning, valued for their simplicity and interpretability. Quinlan's introduction of the C4.5 algorithm in 1993 marked a significant milestone in their development. These models are commonly used in classification tasks, feature selection, and data exploration due to their intuitive approach, which closely resembles human decision-making processes.[23]

### Support Vector Machines (SVM)

Support Vector Machines (SVMs) are a type of supervised machine learning technique that centers around the concept of a "margin," which is the space on either side of a hyperplane that divides two data classes. The primary objective of SVMs is to maximize this margin, striving to create the widest possible separation between the hyperplane and the data points on both sides. This margin maximization has been shown to reduce the upper bound on the expected generalization error, improving the model's performance on unseen data.[24]

In the case of linearly separable data, once the optimal separating hyperplane is established, the data points that lie directly on the margin are referred to as support vectors. The solution is expressed as a linear combination of only these support vectors, while other data points are disregarded. As a result, the model complexity of an SVM does not depend on the total number of features present in the training data. Typically, the SVM learning algorithm selects a relatively small subset of support vectors, making SVMs particularly effective for tasks where the number of features significantly exceeds the number of training instances.

Although the maximum margin concept enables Support Vector Machines (SVMs) to select from multiple candidate hyperplanes, certain datasets may present challenges where no separating hyperplane can be found. This typically occurs when the data contains misclassified instances. To overcome this limitation, the soft margin approach is used, allowing for some level of misclassification among the training instances. This approach balances maximizing the margin while tolerating a few errors to improve the model's flexibility.[25]

In the context of this paper, an exhaustive grid search was conducted to fine-tune the SVM model's hyperparameters. To identify the most suitable Support Vector Machine

(SVM) model for the classification task, three distinct strategies were employed: one vs one, one vs rest, and all vs all of which one vs rest provided the highest accuracy. These approaches aimed to leverage the SVM's ability to handle multi-class classification scenarios effectively.

## METHODOLOGY

This paper introduces three distinct frameworks for classifying surfaces, each utilizing point cloud data in a unique manner. Feature extraction will be performed using three different methodologies. The first approach will employ the MPCA technique, which efficiently captures the multidimensional nature of the point cloud data—represented as a tensor with its duplicate -in a more compact form, thereby enhancing the subsequent data processing tasks and applications. The second approach focuses on extracting variogram's range and sill features, while the third approach will classify surface roughness using $S_a$ surface roughness values.

### DATA PREPROCESSING

Point cloud data can be represented as matrices, where surface heights are the response variables. Point cloud data was collected from three distinct unique surfaces of the C-9 comparator. To improve computational efficiency and Sample size, the regularly gridded point cloud data are segmented. For the MPCA framework, assume pre-processed (after segmenting) point cloud matrices with a Sample size of n1 × n2. A dataset with $n3$ part surface observations can be represented as a tensor object with $n3$ layers of point cloud and $n3$ layers of duplicate point clouds. The output of this step is a third-order tensor denoted as $\boldsymbol{\mathcal{A}} \in \boldsymbol{R^{I_1 \times I_2 \times I_3}}$ where $\mathcal{A}_i \in \boldsymbol{R^{n_1 \times n_2 \times 2}}$ is a tensor for the $i^{th}$ observed part surface. This third-order tensor data can be used directly in the MPCA framework. Segmented point cloud data without tensor formation are used for the variogram features and $Sa$.

### FEATURE EXTRACTION AND DIMENSIONALITY REDUCTION VIA MPCA

MPCA on 3rd Order Tensor proposed by Lu.et.al. (2008) is used in this analysis. For our tensor $\boldsymbol{\mathcal{A}} \in \boldsymbol{R^{I_1 \times I_2 \times I_3}}$, we start by centring the data. The mean tensor $\bar{\mathcal{A}}$ is computed from the entire dataset using Eqn. 2, and each Sample tensor is adjusted using Eqn. 3

$$\bar{\mathcal{A}} = \frac{1}{n3} \sum_{i=1}^{n3} \mathcal{A}_i \qquad \textbf{Eqn. 2}$$

$$\widehat{\mathcal{A}_i} = \mathcal{A}_i - \bar{\mathcal{A}} \qquad \textbf{Eqn. 3}$$

Using high-order singular value decomposition (HOSVD) a 3rd- order tensor A can be decomposed as shown in Eqn. (4).

$$\mathcal{A} = \mathrm{S} \times \mathrm{U}^{(1)'} \times \mathrm{U}^{(2)'} \qquad \textbf{Eqn. 4}$$

where $S \in \mathrm{R}^{n1 \times n2 \times 2n3}$ is a core tensor and $U^{(1)} \in \mathrm{R}^{n1 \times n1}$ and $U^{(2)} \in \mathrm{R}^{n2 \times n2}$ are the orthogonal unitary matrices. The original 3rd-order tensor A can be projected to a target lower dimension tensor subspace $R^{p1} \otimes R^{p2}$( $\otimes$ is the Kronecker product and $p1 < n1,\ p2 < n2$ ). MPCA aims to find two optimal subspaces $U^{(1)}$and $U^{(2)}$that maximize the variance in tensor subspace as shown in Eqn. 5.

$$\text{Max: } \sum_{i=1}^{n3} \left\| (\mathcal{A}_i - \bar{\mathcal{A}}) \times \widehat{U}^{(1)'} \times \widehat{U}^{(2)'} \right\|_F^2 \qquad \textbf{Eqn. 5}$$

$$Subject\ to: rank\ (\widehat{U}^{(1)'}) < n_1\ and\ rank\ \widehat{U}^{(2)'}) < n_2$$

The Frobenius norm, $\|\,.\,\|_F$ is defined as $\|Z_F\| = \sqrt{Z, Z}$, where Z,Z is the inner product of two tensors. Although no closed-form solution exists for the above optimization problem, a solution can be iteratively determined by addressing $\{\widehat{\mathrm{U}}^{(n)} \in \boldsymbol{R^{p_1 \times p_2}}\ \mathrm{n} = 1,2.\}$.

Post the optimization step to derive the multilinear transformation matrices, the transformed features for each observation *I* to a reduced dimension (in this case, a 2nd-order matrix) represented as $y_i$ can be expressed in Eqn. 6:

$$y_i = \widehat{\mathcal{A}_i} \times \widehat{U}^{(1)'} \times \widehat{U}^{(2)'} \qquad \textbf{Eqn. 6}$$

This adaptation of MPCA to our 3rd order tensor paves the way for efficient and robust analysis of combined point cloud and its duplicate datasets.

### FEATURE EXTRACTION USING VARIOGRAM

In this framework, surface roughness features are extracted by analyzing the spatial dependence of point cloud data using variograms. The empirical variogram is first developed in an omnidirectional manner to capture the spatial variance of surface heights across different directions. Examples of omnidirectional variogram for our dataset are shown in Figure 1 . The empirical variogram $\gamma(h)$ is defined in Eqn. 7:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_{i+h})]^2 \qquad \textbf{Eqn. 7}$$

where $\gamma(h)$represents the lag distance between spatial locations, and $(N(h))$ is the number of point pairs separated by the distance ( $h$ ).

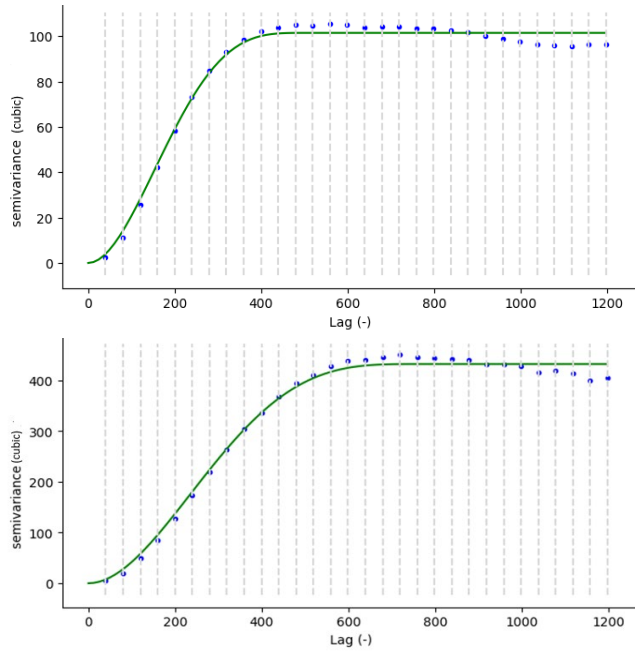



***Figure 1. Omnidirectional variogram (Semivariance 0-400).***

The C-9 comparator surfaces are assumed to be isotropic, which was initially confirmed by constructing both horizontal and vertical empirical variograms that exhibited similar patterns. Based on this observation, the decision was made to construct omnidirectional variograms for further analysis. After constructing the omnidirectional empirical variogram, a theoretical variogram model is fitted by iteratively exploring various models, including spherical, cubic, Matern, Gaussian, and exponential variograms. The best-fitting model is selected by minimizing the Maximum Likelihood Estimation (MLE) score, where the MLE is calculated using the formula in Eqn. 8:

$$\text{MLE} = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}[z(x_i) - m(x_i)]^2$$

**Eqn. 8**

Here, $n$ is the number of data points, $\sigma^2$ is the variance, $z(x_i)$ represents the observed data points, and $m(x_i)$ is the mean modeled by the variogram.

The theoretical variogram that provides the lowest MLE score is selected, and its corresponding range ($r$) and sill ($s$) values are extracted. Most of the empirical variograms fitted with the cubic variogram and to keep the features extracted from the variogram consistent, cubic was chosen to be fitted on all the empirical variograms. The range, which is the distance at which the variogram levels off, indicates the limit of spatial correlation. The sill represents the total variance at this range. These extracted features, [$r$, $s$] are then used as key indicators in the classification of surface roughness, offering a quantitative measure of the spatial structure within the point cloud data.

## FEATURE EXTRACTION USING $S_a$ SURFACE ROUGHNESS VALUES

Surface roughness can also be quantified using the arithmetical mean height, $S_a$, which represents the average deviation of the surface heights from the mean plane across a defined area. The $S_a$ value is calculated in Eqn. 9 as:

$$S_a = \frac{1}{A}\sum_{x,y}|z(x,y) - \bar{z}|$$

**Eqn. 9**

where ( $A$ ) denotes the total area of the surface,$(z(x,y))$is the height at each point $(x,y)$, and $(\bar{z})$ is the mean height across the surface.

In this approach, the $S_a$ value is computed for each segmented area of the point cloud data. The process begins by determining the mean height $(\bar{z})$ of the surface and then calculating the absolute deviations of the heights $(z(x,y))$ from this mean. The mean of these absolute deviations over the area yields the $S_a$ value, which is used as a feature in the classification model. The $S_a$ value effectively captures the overall roughness of the surface, providing a straightforward yet powerful metric for distinguishing between different levels of surface roughness. The $S_a$ data is directly captured for each segment using the preprocessing technique of the 3-D macroscope software system.

## SURFACE QUALITY CLASSIFICATION USING POINT CLOUD DATA

The objective of this part was to enhance the accuracy of surface quality classification by utilizing point cloud data through three distinct feature extraction frameworks. The workflow followed a systematic approach, designed to capture the unique geometric characteristics of the surface from the point cloud data.

The point cloud data was initially segmented into 36 equal parts to ensure consistent analysis across the different feature extraction techniques. This segmentation allowed for a more manageable computational load and ensured that each segment of the surface was uniformly represented.

In the first framework, Multilinear Principal Component Analysis (MPCA) was employed to address the high dimensionality of the tensor data derived from the point clouds. MPCA extracted the most significant features by maximizing the variance within the tensor subspace, reducing the data's dimensionality while preserving the essential structure and relationships. These reduced-dimension features were then input into a machine learning model designed to classify different categories of surfaces roughness based on the geometric information captured in the point cloud data.

The second framework focused on the spatial characteristics of the surface by developing an omnidirectional empirical variogram from the point cloud data. The variogram represented the spatial variance of surface heights across different directions. A theoretical variogram model was then fitted by iterating through spherical, cubic, Matern, Gaussian, and exponential models, with the selection based on the lowest Maximum Likelihood Estimation (MLE) score. The range and sill values from the best-fitting model were extracted as features, providing a quantitative measure of the spatial structure within the point cloud data, which was crucial for surface quality classification.

The third framework utilized the $Sa$ surface roughness values, a standard metric that represents the average height deviation from the mean surface plane. For each segmented point cloud area, the $S_a$ value was obtained by preprocessing the data from the 3-D macroscope for each segment. This straightforward metric captured the overall roughness of the surface, serving as an effective feature for distinguishing between different surface conditions.

Throughout the testing phase, the model's performance was evaluated using the extracted features from the point cloud data alone. This approach provided a consistent data structure, ensuring that the model could accurately classify surface roughness based solely on the geometric information contained within the point clouds.

This research highlights the efficacy of using point cloud data exclusively for surface quality classification. By employing diverse feature extraction techniques—MPCA, variogram analysis, and $Sa$ values—the study ensures a comprehensive and reliable classification process, demonstrating the robustness of the proposed methodologies.

## CLASSIFICATION MODELING

For our surface classification, point clouds are the only data modality used. Merging point cloud with their duplicate, a tensor representation was formulated that captured the geometric intricacies denoted by $\boldsymbol{\mathcal{A}_{train}} \in \boldsymbol{R}^{n_1 \times n_2 \times 2M}$, where M is the historically observed roughness surfaces. Then as explained earlier in the MPCA section, $\boldsymbol{\mathcal{A}_{train}}$ is projected to a lower dimension space, meaning two mapping matrices $\widehat{U}^{(1)'}_{train}$ and $\widehat{U}^{(2)'}_{train}$ that maximize the total scatter in $\boldsymbol{\mathcal{A}_{train}}$ need to be determined. This can be obtained using the following equation:

$$y_{train} = y_i = \mathcal{A}_{train} - \overline{\mathcal{A}_{train}} \times \widehat{U}^{(1)'} \times \widehat{U}^{(2)'} \quad (9)$$

$$\overline{\mathcal{A}_{train}} = \frac{1}{M}\sum_{i=1}^{M} \mathcal{A}_{train,i} \quad (10)$$

The refined data after MPCA was fed into a machine learning model. Notably, the testing phase utilized the tensor structure, with each point cloud being twinned, ensuring consistent data structure for accurate evaluation**.**

While extracting features using variogram framework, the range is found by the distance at which the variogram reaches its sill or levels off as beyond this, the datapoints are not spatially correlated. The sill is the value the variogram reaches at the range. The range and sill formula for each segment depended on the best fitting theoretical variogram. Most of the segments' empirical variogram fitted well with cubic variogram and the remaining fitted with gaussian variogram. The cubic and gaussian variogram models are expressed below:

*Cubic Variogram:*

$$\gamma(h) = c_0 + c\left[7\left(\frac{h}{a}\right)^2 - 8.75\left(\frac{h}{a}\right)^3 + 3.5\left(\frac{h}{a}\right)^5 - 0.75\left(\frac{h}{a}\right)^7\right] \text{ for } 0 \leq h \leq a, \text{ and } \gamma(h) = c_0 + c \text{ for } h > a$$

**Eqn. 10**

*Gaussian Variogram:*

$$\gamma(h) = c_0 + c\left[1 - \exp\left(-\frac{h^2}{a^2}\right)\right]$$

**Eqn. 11**

where $\gamma(h)$ is the lag distance, $a$ is the range, and $c_0 + c$ is the sill. Both the range and sill features for all the point cloud dataset were extracted and split into training and testing dataset for classification modeling.

The $S_a$ features which were directly extracted while preprocessing the collected data from the 3-D macroscope for each segment are used for training and testing. This feature was then used to build and apply the classification model using machine learning.

The classification of the methodology can be implemented using an array of algorithms, including decision trees, neural networks, and K-nearest neighbors. When employing standard supervised learning techniques, it is essential to select the optimal performance metrics, like accuracy. When evaluating the efficacy of the classification model using the MPCA feature extraction technique, performance metrics should exclusively focus on the point cloud data structured as tensors. The point cloud data for the remaining two frameworks, Variogram Analysis (Range and Sill) and $S_a$ Surface Roughness Values, are processed separately according to their respective methodologies. To gauge the effectiveness of the proposed method against other benchmark models, a multi-class confusion matrix is utilized.

Table 1 illustrates the multiclass classification metrics and formulas. This matrix represents the actual vs. predicted classifications for every class. Key metrics that are calculated for each class include True positive (TP), True negative (TN), False positive (FP), and False negative (FN). As an illustration, TP represents the count of images from a

specific class that were accurately classified. This model's performance is assessed using prominent multi-class metrics, namely Accuracy and F-score. These metric formulas are derived from the work of Sokolova and Lapalme.[26]

***Table 1. Multi-class Classification Metrics and Formulas***

| Metric | Formula |
|---|---|
| $Accuracy_i^*$ | $\frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i}$ |
| $Accuracy_M$ | $\frac{\sum_{i=1}^{k} \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i}}{k}$ |
| $Precision_\mu$ | $\frac{\sum_{i=1}^{k} TP_i}{\sum_{i=1}^{k} TP_i + FP_i}$ |
| $Recall_\mu$ | $\frac{\sum_{i=1}^{k} TP_i}{\sum_{i=1}^{k} TP_i + FN_i}$ |
| $F_{score_\mu}$ | $\frac{2 \times Precision_\mu \times Recall_\mu}{Precision_\mu + Recall_\mu}$ |
| $Precision_M$ | $\frac{\frac{\sum_{i=1}^{k} TP_i}{\sum_{i=1}^{k} TP_i + FP_i}}{k}$ |
| $Recall_M$ | $\frac{\frac{\sum_{i=1}^{k} TP_i}{\sum_{i=1}^{k} TP_i + FN_i}}{k}$ |
| $F_{score_M}$ | $\frac{2 \times Precision_M \times Recall_M}{Precision_M + Recall_M}$ |

## CASE STUDY

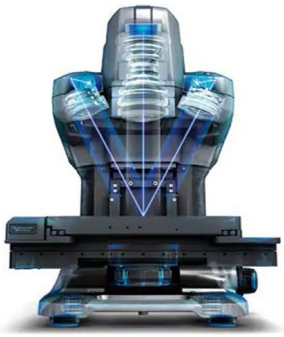
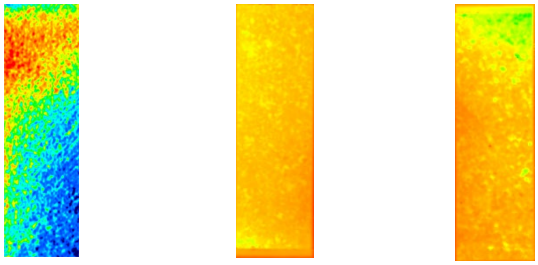


***Figure 3. Schematic of one-shot 3D measuring macroscope.***

The effectiveness of our model was explored through our case study of surface classification of C-9 comparator surfaces. In this study, point cloud data and the $S_a$ values from the comparator were obtained. The 3 unique comparator surfaces' point cloud images are shown in Fig. 3. This section provides a complete description of the experimental setup, data collection and the results obtained.



(a) (b) (b)

***Figure 2. Unique surfaces of the C-9 comparator of (a) 200 µinches; (b) 300 µinches; (c) 560 µinches.***

## EXPERIMENTAL SETUP

To achieve a high level of consistency in surface classification, this case study utilized 3 unique surfaces of the compared, valued for their uniform textural qualities. The selection of these surfaces allowed for clear differentiation across various roughness categories and consistent data collection within each class, crucial for the development and validation of our classification model.

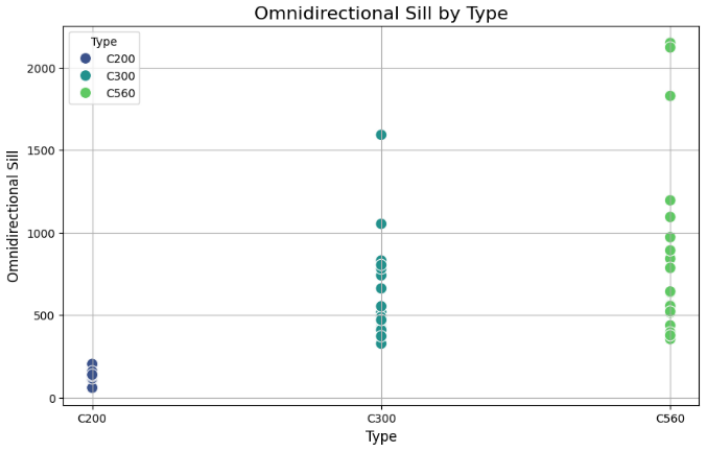



***Figure 4. Scatter plot of type versus omnidirectional sill.***

### Data Collection

Data for each comparator surface was captured using a KEYENCE VR-3200 one-shot 3D measuring macroscope (seen in Fig. 3). This advanced instrument captures the surface's topography in the form of structured point clouds in a short period of time (approximately 4 s). It operates as a high-resolution imaging system capable of capturing intricate details on surfaces. The macroscope simultaneously captures both digital images and point clouds both of which had a 2-D dimensions of approximately 36000 x 9000µm. As this research only requires point cloud data, these observations were further segmented to 36 parts of dimension 3000 x 3000 mm. The macroscope also helps in surface roughness analysis where it analyzes each segment and quantifies the surface roughness using multiple surface roughness metrics of which '$S_a$' is the metric we consider. A

high pass filter of 0.25mm was applied to each segment before the $S_a$ metric was collected.

### DATA PRE-PROCESSING

The primary step in our data pre-processing routine involved segmenting the data, given the vast expanse of the comparator unique surface specimen which ensured the captured data was solely from within the specimen. As the original image dimension was of the point cloud dimension was 36000 x 9000µm, it was segmented into thirty-six uniform partitions, each measuring 3000 x 3000µm. After segmentation, median polish was applied to each segment of the point cloud data to remove any systematic surface trends which may influence the classification process. This segmentation approach resulted in a total of 36 observations of point cloud information for each of the three-comparator surfaces resulting in a total of 108 observations. Segmentation helps in increasing the sample size for the analysis.

## RESULTS AND DISCUSSION

In this section, we present the performance evaluation of three distinct frameworks for surface quality classification using point cloud data: MPCA, variogram analysis (range and sill), and $S_a$ surface roughness metric. The aim is to compare the effectiveness of these frameworks in capturing surface roughness characteristics from point cloud data, providing insights into their individual strengths and limitations.

In our research, we developed three models corresponding to each framework. The MPCA model employed Multilinear Principal Component Analysis to reduce the dimensionality of the high-dimensional point cloud data while preserving significant spatial features. The variogram-based model extracted spatial features by fitting theoretical variogram models and calculating range and sill values, and the $S_a$ -based model used the $S_a$ surface roughness metric as the feature for classification.

Our process for model development and validation followed two stages, applied to the full point cloud dataset. The dataset was processed uniformly across all frameworks without dividing into training and testing sets. This approach ensured consistency in the comparison of the frameworks' classification performance.

During the first stage, we employed MPCA to reduce dimensionality, ensuring that critical features indicative of surface roughness was preserved. This was particularly important for extracting and representing 95% of the total variance in the data, allowing the model to focus on the most influential factors in the classification process. In parallel, the variogram-based model developed empirical variograms and fitted them to theoretical models, selecting the best-fitting models based on the lowest Maximum Likelihood Estimation (MLE) score. The range and sill values from the fitted models were used as key features for surface classification. For the $S_a$ -based model, the $S_a$ values, which represent the average height deviation from the mean plane, were calculated for each point cloud segment and used as the primary feature.

The second stage entailed evaluating the models to simulate how each framework would perform when applied to unseen data. The classification accuracy and F1-score per class were recorded for each model, and the performance of each framework was compared.

***Table 2. Accuracy (%) per Label***

| Surface type | $MPCA$ | $Variogram$ | $Sa\ metric$ |
|---|---|---|---|
| C200 | 100 | 100 | 100 |
| C300 | 83.36 | 78.33 | 66.66 |
| C560 | 84.66 | 75 | 76.66 |
| Macro-average accuracy | 89.34 | 84.45 | 81.10 |

***Table 3. $F_{score}$ per Label***

| Surface roughness | $MPCA$ | $Variogram$ | $Sa\ metric$ |
|---|---|---|---|
| C200 | 1 | 1 | 1 |
| C300 | 0.80 | 0.7033 | 0.33 |
| C560 | 0.86 | 0.6918 | 0.5 |
| Macro-average $F_{score}$ | 0.84 | 0.6861 | 0.61 |

From the accuracy across labels, it is evident that surface type C200 has 100% classification accuracy across all three frameworks. For the variogram model, it was evident from the scatter plot of type vs sill that the sill values for C200 type can be completely classified as compared to C300 and C560. This indicates that the sill values have more influence on classification for surface quality than the range feature of the variogram.

Although classification using variogram features yields good accuracy, it is important to note that the semivariance between data points for any comparator surface analyzed flattens at a radius of approximately 0.007 mm. To put this in perspective, each unique comparator surface measured roughly 36 x 9 mm. This indicates that the variogram features are effective for surface classification only within this small range, which is negligible for practical applications in the foundry industry. The reason why MPCA gives higher accuracy is because the number of important

features is more as compared to the remaining two frameworks.

It is also important to note that as the data interpretability decreases, the accuracy increases as seen from the accuracy results for classification. $S_a$ is a good feature from a design standpoint whereas MPCA is a good feature from a monitoring standpoint. This can be represented in a plot as shown in Fig. 5:

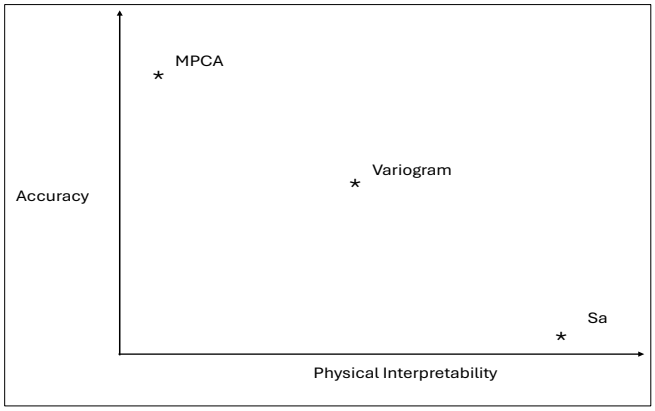



***Figure 5. Physical Interpretability vs accuracy.***

## CONCLUSION AND FUTURE WORK

This study compared three different point cloud feature extraction technique-based models for roughness classification of a flat surface. The investigation showed how the different feature extraction methods classify surface roughness and irrespective of accuracy being good for all of them, it is essential to choose the appropriate technique for classification.

The results demonstrated that the model utilizing MPCA features outperformed the variogram model and the $S_a$ features model in classifying surface roughness. The ability of the $MPCA$ model to capture a more detailed representation of surface texture attributes, owing to the additional depth information provided by point clouds, led to more accurate classification.

It is crucial to emphasize that only point cloud data for 2D surfaces was employed in this study. Future work in this area should concentrate on categorizing a greater variety of surface roughness textures and more geometrically complex surfaces as the field develops. The goal is to create models that can accurately classify data while also effectively predicting surface roughness parameters and adjusting to various surface textures. This would probably entail investigating cutting-edge data processing methods, improving algorithms, and perhaps incorporating different kinds of sensor data to produce a more thorough knowledge of surface properties.

The creation of predictive models for surface quality classification is in line with the overarching objectives of Industry 4.0 and smart manufacturing, as the capacity to anticipate and manage surface quality in real-time can lead to major advancements in production procedures. By improving these models' prediction precision and flexibility, we go one step closer to completely self-governing foundry systems that can self-optimize for surface quality control.

## REFERENCES

1. Whitehouse, D.J. (1994), “Handbook of Surface Metrology,” Institute of Physics Publishing, Bristol, U.K.
2. Bhushan, B. (2000). „Surface roughness analysis and measurement techniques,“ CRC Press eBooks, 79–150. https://doi.org/10.1201/9780849377877-10.
3. Ghosh, G., Kumar, M., Sidpara, A M., & Bandyopadhyay, P. (2022), “Tribological aspects of different machining processes,” Elsevier eBooks (pp. 213–238). https://doi.org/10.1016/b978-0-12-819889-6.00006-x
4. Sithole, G., & Vosselman, G. (2004), Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds, *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(1-2), 85-101.
5. Qi, C.R., Su, H., Mo, K., & Guibas, L.J. (2017). PointNet: Deep learning on point sets for 3D classification and segmentation*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 652-660.
6. Xia, H., Ding, Y., & Mallick, B. K. (2011), Bayesian hierarchisscal model for combining misaligned two-resolution metrology data, *IIE Transactions*, 43(4), 242–258.
7. Suriano, S., Wang, H., Shao, C., Hu, S. J., & Sekhar, P. (2015), Progressive measurement and monitoring for multi-resolution data in surface manufacturing considering spatial and cross correlations*, IIE Transactions*, 47(10), 1033–1052.
8. Helle, R. H., & Lemu, H. G. (2021). A case study on use of 3D scanning for reverse engineering and quality control, *Materials Today: Proceedings*, 45, 5255–5262. https://doi.org/10.1016/j.matpr.2021.01.828.
9. Dastoorian, R., Elhabashy, A. E., Tian, W., Wells, L. J., & Camelio, J. A. (2018, June). Automated surface inspection using 3D point cloud data in manufacturing: A case study*, International manufacturing science and engineering conference* (Vol. 51371, p. V003T02A036). American Society of Mechanical Engineers.
10. Rao, P. K., Kong, Z., Duty, C. E., Smith, R. J., Kunc, V., & Love, L. J. (2016). Assessment of dimensional integrity and spatial defect localization in additive manufacturing using spectral graph theory, *Journal of Manufacturing Science and Engineering,* 138(5).

11. Tootooni, M. S., Liu, C., Roberson, D., Donovan, R., Rao, P. K., Kong, Z. J., & Bukkapatnam, S.T. (2016). Online non-contact surface finish measurement in machining using graph theory-based image analysis, *Journal of Manufacturing Systems*, 41, 266-276.
12. Ding, C., & Li, T. (2007, June). Adaptive dimension reduction using discriminant analysis and k-means clustering, *Proceedings of the 24th international conference on Machine learning* (pp. 521-528).
13. Yang, J., Zhang, D., Frangi, A. F., & Yang, J. Y. (2004). Two-dimensional PCA: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1), 131–137.
14. Jolliffe, I. T. (2002). Principal component analysis for special types of data (pp. 338-372). Springer New York.
15. Colosimo, B. M., & Grasso, M. (2018). Spatially weighted PCA for monitoring video image data with application to additive manufacturing, *Journal of Quality Technology*, 50(4), 391–417.
16. Lu, H., Plataniotis, K. N., & Venetsanopoulos, A. (2008). MPCA: Multilinear Principal Component Analysis of Tensor Objects, *IEEE Transactions on Neural Networks*, 19(1), 18–39. https://doi.org/10.1109/tnn.2007.901277.
17. Shetty, R., Al Majali, A., & Wells, L. (n.d.). Surface Insight: Leveraging High-Density Dataset Fusion for Enhanced Roughness Classification, *NAMRC*. 52nd SME North American Manufacturing Research Conference (NAMRC 52, 2024), United States of America.
18. Schimpf, D. W., & Peters, F. E. (2021). 3D analysis of casting surface characterization based on the Variogram Roughness Method. *International Journal of Metalcasting*, *16*(3), 1079-1090, https://doi.org/10.1007/s40962-021-00693-6.
19. Chen, H. (2022). Techniques of fine reservoir description, *Elsevier eBooks* (pp. 113–348). https://doi.org/10.1016/b978-0-323-95401-3.00003-2.20.
20. Schimpf, D. W., & Peters, F. E. (2020). Variogram roughness method for casting surface characterization, *International Journal of Metalcasting*, *15*(1), 17–28. https://doi.org/10.1007/s40962-020-00451-0
21. Ding, C., & Li, T. (2007, June). Adaptive dimension reduction using discriminant analysis and k-means clustering, *Proceedings of the 24th international conference on Machine learning* (pp. 521-528).
22. Shetty, R., Majali, A. A., & Wells, L. (2024). Enhanced Classification of refractory coatings in foundries: a VPCA-Based Machine learning approach. *International Journal of Metalcasting*. https://doi.org/10.1007/s40962-024-01427-0
23. Quinlan, J. R. (1992). C4.5: Programs for Machine learning. https://cds.cern.ch/record/2031749
24. Castelli, M., Vanneschi, L., & Largo, Á. R. (2019). Supervised Learning: classification, *Elsevier eBooks* (pp. 342–349). https://doi.org/10.1016/b978-0-12-809633-8.20332-4
25. Veropoulos K, Campbell C, Cristianini N (1999) Controlling the sensitivity of support vector machines. *Proceedings of the international joint conference on artificial intelligence* (IJCAI99).
26. Sokolova, M., & Lapalme, G. (2009), A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4), 427-437.