
Proceedings Paper:
Papananias, M. orcid.org/0000-0001-7121-9681, Obajemu, O., McLeay, T.E. et al. (2 more authors) (2020) Development of a new machine learning-based informatics system for product health monitoring. In: Gao, R.X. and Ehmann, K., (eds.) Procedia CIRP. 53rd CIRP Conference on Manufacturing Systems, 01-03 Jul 2020, Chicago, IL, U.S.. Elsevier , pp. 473-478.

https://doi.org/10.1016/j.procir.2020.03.075

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 93 (2020) 473–478

www.elsevier.com/locate/procedia

53rd CIRP Conference on Manufacturing Systems

Development of a New Machine Learning-based Informatics System for Product Health Monitoring

Moschos Papananias[a]*, Olusayo Obajemu[a], Thomas E McLeay[b], Mahdi Mahfouf[a], Visakan Kadirkamanathan[a]

[a]*Department of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK*
[b]*Sandvik Coromant, Mossvägen 10, Sandviken 811 34, Sweden*

* Corresponding author.*E-mail address:* m.papananias@sheffield.ac.uk

Abstract

Manufacturing informatics aims to optimize productivity by extracting information from numerous data sources and making decisions based on that information about the process and the parts being produced. Manufacturing processes usually include a series of costly operations such as heat treatment, machining, and inspection to produce high-quality parts. However, performing costly operations when the product conformance to specifications cannot be achievable is not desirable. This paper develops a new machine learning-based informatics system capable of predicting the end product quality so that non-value-adding operations such as inspection can be minimized and the process can be stopped before completion when the part being manufactured fails to meet the design specifications.

Keywords: Manufacturing Informatics; Multistage Manufacturing Process; Principal Componet Analysis; Artificial Neural Networks; Multiple linear Regression

1. Introduction

Smart manufacturing is the application of machine learning in manufacturing processes to improve process performance and product quality, reduce scrap and non-conformance costs, and minimize environmental impacts [1]. Machine learning is concerned with the development of software systems that can extract knowledge from data, produced usually from a wide range of sources, and determine patterns and make predictions or decisions on the basis of these uncovered patterns. Machine learning algorithms are based on statistical and soft computing techniques that allow them adaptively improve their performance as more data become available for learning and make decisions without human intervention. These algorithms are divided into two major categories, depending on their learning mechanism: supervised and unsupervised [2]. Supervised learning algorithms such as Artificial Neural Networks (ANNs) learn a functional model from observed input-output samples to derive relationships and make predictions on outcomes of interest when new input data become available. Unsupervised learning algorithms learn only from input data (unlabeled data) and can be used for various machine learning problems such as clustering and dimensionality reduction [3, 4]. There is an important third category of learning algorithms called semi-supervised, which learn from partially labeled datasets. Semi-supervised learning algorithms are outside the scope of this work, but interested readers are referred to [5].

With the advances in sensor technologies, the amount of sensor data obtained from manufacturing processes has increased explosively [6, 7]. Manufacturing processes usually include various stages such as heat treatment, quenching and

2212-8271
10.1016/j.procir.2020.03.075

tempering, a series of metal-removing operations, and inspection to produce a part. Data-driven approaches for manufacturing process monitoring and control can provide real-time insights into process performance and thus, part quality can be improved by adjusting certain process parameters such as feed rate, spindle speed and depth of cut in machining operations. Coordinate Measuring Systems (CMSs) such as Coordinate Measuring Machines (CMMs) or automated comparator gauges are used in the manufacturing industry for inspection tasks [8]. Such measurement systems gather data points related to the part surface. Then, these data points are processed through software to obtain the quality characteristics of interest and determine whether the manufactured part conforms to design specifications [9, 10].

Building a mathematical model based on measured data with or without a priori knowledge about the system under consideration is commonly encountered in diverse areas of manufacturing including process and product health monitoring and control. In this area, the machine learning problem is often characterized by the presence of a small number of high-dimensional training samples due to the high cost of workpiece materials, cutting tools, etc. Therefore, dimensionality reduction algorithms are usually required to transform the high dimensional dataset into a lower-dimensional space. This reduction can be achieved by various techniques such as Principal Component Analysis (PCA) and ANNs [11, 12]. PCA is a well-established feature extraction method that can be used to perform a linear mapping of the collected process data to a lower-dimensional space [13]. The field of ANNs has made huge progress in the past four decades. ANNs, inspired by the biological neural networks of the human brain, are one of the main computing systems used in machine learning for pattern recognition and function approximation problems [14, 15]. The most popular type of ANNs is the Multi-Layer Perceptron (MLP) network, which is a feedforward neural network. MLP networks can also be used in a self-supervised or auto-associative mode to perform nonlinear dimensionality reduction. Self-supervised MLP networks, also known as autoencoders, are trained in an unsupervised manner since no labeled data are needed, i.e., the input data and the target data are identical.

The subject of monitoring and control of manufacturing processes is a broad and fast-moving one due to recent advances in sensor technology and information processing algorithms as well as, increasing demands for higher productivity, improved product quality and reduced costs [16]. The published research in this area is extensive, particularly for tool wear and surface roughness monitoring applications and the applied methodologies may depend on application and the manufacturing process and material, among other factors such as costs, tolerances, and process variability. Risbood et al. [17] developed neural networks to predict surface roughness and dimensional deviation using force and vibration data in cylindrical turning. Kovač et al. [18] developed fuzzy logic and regression models to predict surface roughness in face milling using cutting speed, feed rate, depth of cut and flank wear land width. García et al. [19] examined feature extraction methods to optimize surface finish monitoring using vibration signals. Arnaiz-González et al. [20] developed and compared neural network models including MLP networks and Radial Basis Functions (RBFs) to predict the dimensional error on inclined surfaces obtained by ball-end milling using various process variables including cutting type and strategy, tool slenderness, material hardness, and surface slope. Gao et al. [21] presented a product quality monitoring system based on Support Vector Regression (SVR) to predict the thickness and width of injection moulded parts using four process variables including melt pressure, temperature, velocity and viscosity. Papananias et al. [22] developed a probabilistic model based on Bayesian linear regression to estimate part quality and associated uncertainties for Multistage Manufacturing Processes (MMPs). Wang et al. [23] applied neural networks for dimensionality reduction and prediction of defective parts produced by a powder metallurgy process. Li et al. [24] described a deep learning-based classification model to detect defective parts using the concept of fog computing in order to process big data in real time. Compared to tool wear and surface quality monitoring, published research on monitoring systems for dimensional metrology characteristics is limited and much of it focuses on artificial intelligence methods based on single-stage manufacturing data. However, manufacturing processes usually involve multiple stages of manufacturing to produce a part and thus, part quality is influenced by a large range of error sources introduced by the current, as well as preceding, manufacturing stages [25].

The purpose of this paper is to develop a monitoring system that learns from multistage manufacturing data and predicts the dimensional metrology characteristics of the parts being produced. A major advantage of the proposed methodology is the ability to reduce the volume of non-value-adding operations such as inspection. Another advantage of the proposed methodology is the possibility to stop the manufacturing process at a certain step before performing additional operations when the product conformance to specifications is not achievable. The proposed methodology also allows one to revise the manufacturing strategy and adjust certain process parameters for re-work before removing the workpiece from the machine as in the case of post-process inspection. The methodology is validated using experimental data with aim to illustrate the ability of the proposed methodology to generalize across different geometric parameters of the workpiece including diameter, true position (2-dimensional), and circularity (2-dimensional). The true position tolerance specifies how far the location of a feature can deviate from its "true position" using a 2-dimensional or 3-dimensional tolerance zone. Dimensional metrology characteristics such as circularity are important form tolerances applied to individual features. A form tolerance is used to control the permissible variation of a feature from a perfect geometric element [8].

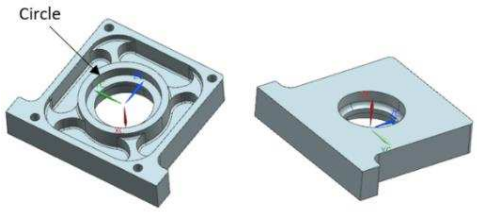

Fig. 1. CAD model of the bearing housing part.

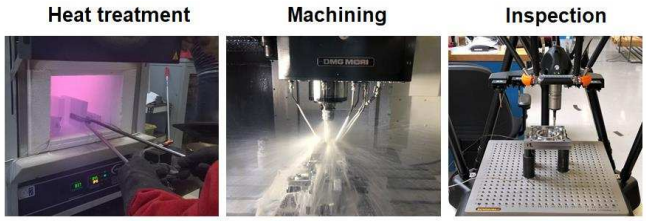

Fig. 2. Heat treatment, machining and inspection processes.

Section 2 describes the experimental work performed to obtain metrology data from different sources and stages of manufacturing. Sections 3 and 4 present the proposed methodology and results, respectively. Concluding remarks and suggestions for further work are given in section 5.

2. Experimental work

The solutions for process monitoring and control systems for manufacturing vary with the application areas and type and quantity of measurement data. Therefore, a specific manufacturing case study has been developed, involving multiple stages of manufacturing. Experimental work was performed to gather metrology data as each product went through the steps of heat treatment, machining, and inspection. A VECSTAR furnace was used to heat treat the material blocks. In particular, the material blocks were heated up to 845°C and then quenched in oil for hardening. To add variability in material properties, the blocks were tempered at different temperatures including 450°C, 550°C, and 650°C. The temperature of the furnace was measured using five K-type thermocouples. The heat treated material blocks were then grinded to improve the surface quality and measure the material surface hardness using a Rockwell device. A DMG MORI NVX 5080 3-axis Computer Numerically Controlled (CNC) machine tool was used for machining, where a full factorial design was performed with four factors at two levels and one center point each so in total seventeen experimental runs were required. The factors were: material surface hardness, feed rate, spindle speed, and datum error in both X and Y axes when the workpiece was flipped around the Y axis for machining the features of the bottom side of the workpiece. Therefore, top side features such as the circle (see Fig. 1) are subject only to the first three factors. Before machining each workpiece, all the cutting tools were evaluated for wear on each flute using a Leica microscope and changed when reached a certain flank wear width. Vibration data were obtained at 10 kHz using an accelerometer sensor, placed on the spindle, and NI LabVIEW SignalExpress software.

For post-process inspection, the Renishaw Equator 300 Extended Height gauging system was employed on the shop floor using the SP25 scanning probe, which can function either as a scanning or touch-trigger probe [26]. The Equator is a comparator measurement system based on parallel kinematic structure. The Equator was used to perform both discrete-point probing and scanning comparator measurements using the CMM Compare method, which requires generating a calibration file for the comparator system by measuring a production part produced close to drawing nominals on a calibrated CMM [27]. The CMM used to generate the calibration file was a Mitutoyo CMM with a Renishaw REVO RSP3 3D scanning probe. The Mitutoyo CMM was located in stable temperature controlled conditions because temperature changes affect the scales, machine, and parts being measured. A typical 30 mm long stylus with tungsten carbide stem and a 2 mm diameter ruby ball was used for both CMSs. The production part labelled as ‘master’ part had been thermally stabilized before starting the CMM measurement task. The Equator inspection approach is based on a point-to-point comparison between the master part measurement data and the test part measurement data. Therefore, special attention was paid to the fixturing setup to minimize the influence of part misalignment from rotation between master and measure coordinate frames on the uncertainty associated with comparative coordinate measurement [14]. The same part fixturing setup was used for both measurement processes using modular fixturing components including magnets and pin magnets. Fig. 1 shows the Computer-Aided Design (CAD) model of the part. Fig. 2 shows the main production processes for the manufacture of the steel bearing housing parts. Fig. 3 shows the proposed framework of machine learning-based informatics systems for product health monitoring.

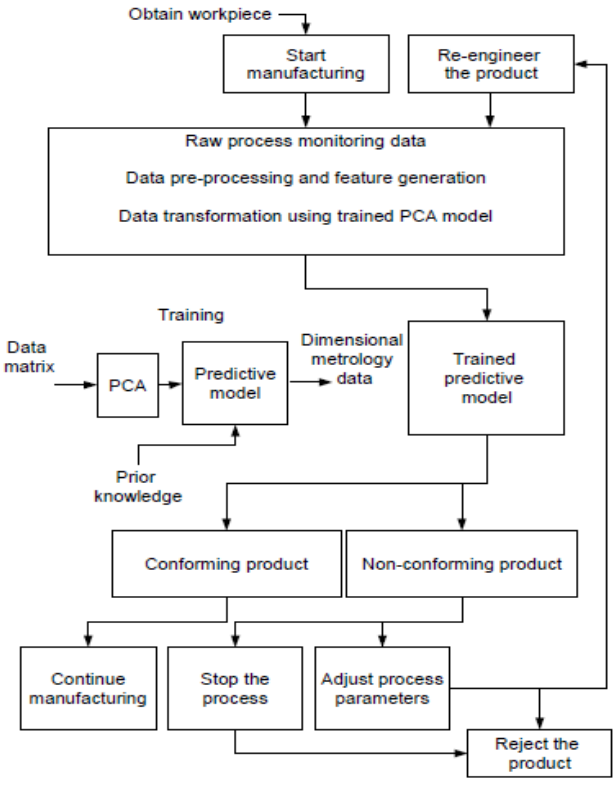

Fig. 3. Proposed framework.

3. Proposed machine learning-based informatics system

This section presents the proposed machine-learning based informatics system for monitoring the dimensional metrology characteristics of the parts being manufactured. Manufacturing informatics is a relatively new area that concerns the monitoring and control of production processes and product variability using live captured sensor data. Our approach utilizes machine learning algorithms including both unsupervised and supervised algorithms to map the extracted process features to the geometric parameters of the workpiece.

The input data in this work included: a) the maximum temperature of the furnace obtained from the tempering process; b) a three-state variable for the material surface hardness measurements; and c) time-domain features including Root-Mean-Square (RMS), sample kurtosis, sample skewness, sample variance and mean of vibration components Vx, Vy, and Vz. The input data were normalized by the Euclidean norm (2-norm). The output data included the product quality deviations obtained from the Equator scanning measurements using the CMM Compare method. The deviations were calculated by the difference between the drawing nominal value and the measured value. To reduce the dimensionality of the input data from q to $n(\ll q)$, PCA was performed using a Singular Value Decomposition (SVD) of the normalized input data matrix, $\boldsymbol{D} \in \mathcal{R}^{m\times q}$:

$$\boldsymbol{D} = \boldsymbol{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}, \tag{1}$$

where $\boldsymbol{U} \in \mathcal{R}^{m\times m}$ and $\mathbf{V} \in \mathcal{R}^{q\times q}$ are orthogonal matrices and $\boldsymbol{\Sigma} \in \mathcal{R}^{m\times q}$ is a diagonal matrix containing diagonal elements $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,q)} \geq 0$.

A linear regression model was developed for each dimensional quality characteristic of interest using as inputs the first four principal components extracted from the normalized input data matrix. A linear regression model can be defined by:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{a} + \boldsymbol{\epsilon}, \tag{2}$$

where $\boldsymbol{y} \in \boldsymbol{Y}$ is an $m \times 1$ observation vector with $E(\boldsymbol{Y}) = \boldsymbol{X}\boldsymbol{a}$, $\boldsymbol{\epsilon} \in \boldsymbol{\mathcal{E}}$ is an $m \times 1$ error vector with $E(\boldsymbol{\mathcal{E}}) = \mathbf{0}$ and $V(\boldsymbol{\mathcal{E}}) = V(\boldsymbol{Y}) = E(\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{E}}^{\mathrm{T}}) = \sigma^2\boldsymbol{I}_m$, $\boldsymbol{X} = [\mathbf{1}_m\, \boldsymbol{x}_1 \ldots \boldsymbol{x}_n]$ is an $m \times (n+1)$ matrix of covariates, $\boldsymbol{a}$ is an $(n+1)\times 1$ vector of unknown parameters, σ^2 is the unknown variance parameter, and $\boldsymbol{I}_m$ is the $m \times m$ identity matrix with ones on the diagonal and zeros elsewhere. If $\boldsymbol{X}$ is full rank, so that $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ is nonsingular, the least-squares estimator $\boldsymbol{a}$ of $\boldsymbol{a}$ is:

$$LSE(\boldsymbol{a}) = \boldsymbol{a} = \boldsymbol{X}^{\dagger}\boldsymbol{y}, \tag{3}$$

where $\boldsymbol{X}^{\dagger} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$ denotes the Moore-Penrose inverse of $\boldsymbol{X}$. If $\boldsymbol{X}$ is not full rank, $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$ is replaced by a generalized inverse $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-}$ of $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$. Assuming that $E(\boldsymbol{\mathcal{E}}) = \mathbf{0}$ and $V(\boldsymbol{\mathcal{E}}) = \sigma^2\boldsymbol{I}_m$, then the least-squares estimator $\boldsymbol{a} \in \boldsymbol{A}$ is an unbiased estimator of $\boldsymbol{a}$, $E(\boldsymbol{A}) = \boldsymbol{a}$, and $V(\boldsymbol{A}) = \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$. If the vector of random variables $\boldsymbol{\mathcal{E}}$ follows a multivariate normal distribution with expectation $\mathbf{0}$ and covariance matrix $\sigma^2\boldsymbol{I}_m$, $\boldsymbol{\mathcal{E}}\sim \mathrm{N}_m(\mathbf{0}, \sigma^2\boldsymbol{I}_m)$ then the least-squares estimator $\boldsymbol{a} = LSE(\boldsymbol{a})$ is also the maximum likelihood estimator $MLE(\boldsymbol{a})$ of $\boldsymbol{a}$, $LSE(\boldsymbol{a}) = MLE(\boldsymbol{a})$, and the vectors of random variables $\boldsymbol{Y}$ and $\boldsymbol{A}$ are also normally distributed, $\boldsymbol{Y}\sim \mathrm{N}_m(\boldsymbol{X}\boldsymbol{a}, \sigma^2\boldsymbol{I}_m)$ and $\boldsymbol{A}\sim \mathrm{N}_{n+1}(\boldsymbol{a}, \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1})$, respectively. An estimate of σ^2 is given by:

$$\hat{\sigma}^2 = \frac{\hat{\boldsymbol{\epsilon}}^{\mathrm{T}}\hat{\boldsymbol{\epsilon}}}{m-n-1} = \frac{(\boldsymbol{y}-\hat{\boldsymbol{y}})^{\mathrm{T}}(\boldsymbol{y}-\hat{\boldsymbol{y}})}{m-n-1}, \tag{4}$$

where $\hat{\boldsymbol{\epsilon}} = \boldsymbol{y} - \hat{\boldsymbol{y}}$ is the vector of residuals and $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{a}$ denotes the fitted values of $\boldsymbol{y}$. Therefore, we can estimate the covariance matrix of $\boldsymbol{A}$ by:

$$\hat{V}(\boldsymbol{A}) = \hat{\sigma}^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1} = \frac{\hat{\boldsymbol{\epsilon}}^{\mathrm{T}}\hat{\boldsymbol{\epsilon}}}{m-n-1}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}. \tag{5}$$

For numerical accuracy and computational efficiency the $\boldsymbol{QR}$ factorization, $\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{R}$, is computed, where $\boldsymbol{Q}$ is an $m \times (n+1)$ matrix of orthonormal columns ($\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{I}_{n+1}$) and $\boldsymbol{R}$ is an $(n+1)\times(n+1)$ upper triangular matrix [28, 29].

In addition, an MLP network with one hidden layer and ten hidden neurons was developed for each dimensional quality characteristic using as inputs the first four principal components. Tan-sigmoid transfer functions were used in the hidden and output layers. Each network was trained by conjugate gradient backpropagation with Powell-Beale restarts. The dimension of the hidden layer was determined by a trial and error process. In MLP networks, every node in any layer of the network is connected with a certain weight to every node in the subsequent layer. The nodes in the hidden and output layers are also called artificial neurons because they include a summation unit and an activation function. An artificial neuron computes a weighted sum over its inputs, adds a bias or threshold term to the sum, and produces an output by transforming the sum through a continuously differentiable nonlinear activation function [15]. However, output neurons with linear activation functions are also common since, in machine learning problems such as regression and classification, nonlinearity only in the hidden neurons is required to achieve nonlinear mapping between the input and output data.

4. Results

A 4-fold cross-validation procedure was performed in MATLAB to evaluate the performance of machine learning models on unseen data. In particular, the dataset obtained from the manufacture of seventeen parts was partitioned into four folds, three of which included data from four parts and one fold included data from the remaining five parts. Three folds were used for training and one fold was used for testing. The cross-validation process was repeated four times so that all folds were used as the validation dataset once. Table 1 provides the PCA results including the percent variability explained by the first four principal components during training, the training reconstruction Root Mean Squared Error (RMSE) considering the first four principal components, and the testing reconstruction RMSE.

Table 1. PCA results.

Folds	Sum of variance explained	Training reconstruction RMSE	Testing reconstruction RMSE
Fold 1	96.4677%	0.00964	0.01543
Fold 2	97.1512%	0.00881	0.01670
Fold 3	96.3988%	0.00982	0.01480
Fold 4	97.1110%	0.00919	0.01591
Weighted average	96.8015%	0.00936	0.01572

Table 2. 4-fold cross validations results of neural network and linear models for diameter deviation.

Folds	Neural network		Linear regression	
	Training RMSE (μm)	Testing RMSE (μm)	Training RMSE (μm)	Testing RMSE (μm)
Fold 1	0.68	1.81	3.74	8.12
Fold 2	0.75	2.59	4.93	4.99
Fold 3	0.06	1.40	4.75	5.81
Fold 4	1.07	3.13	3.86	7.58
Weighted average	0.67	2.29	4.29	6.68

Table 3. 4-fold cross validations results of neural network and linear models for circularity.

Folds	Neural network		Linear regression	
	Training RMSE (μm)	Testing RMSE (μm)	Training RMSE (μm)	Testing RMSE (μm)
Fold 1	0.42	2.82	0.98	3.78
Fold 2	0.20	0.76	1.66	2.20
Fold 3	0.76	1.17	1.73	2.50
Fold 4	1.35	1.40	1.76	1.80
Weighted average	0.72	1.53	1.54	2.52

Table 4. 4-fold cross validations results of neural network and linear models for true position.

Folds	Neural network		Linear regression	
	Training RMSE (μm)	Testing RMSE (μm)	Training RMSE (μm)	Testing RMSE (μm)
Fold 1	0.65	1.04	1.00	1.12
Fold 2	1.57	1.11	0.90	1.44
Fold 3	0.32	0.68	0.98	0.98
Fold 4	1.17	1.06	0.70	1.74
Weighted average	0.94	0.98	0.88	1.34

Tables 2-4 show the RMSE results obtained from the linear regression and neural network models. The neural network training epochs were, on average, 183 for diameter deviation, 199 for circularity, and 133 for true position, and, therefore, the training times can be considered to be negligible (less than one second). Based on Tables 2-4, it can be concluded that the neural network models achieved more accurate predictions than the linear regression models for all three measurands of interest. For the neural network model, the residual values range from 0.3 μm to 6.1 μm for diameter deviation, from 0.1 μm to 5.6 μm for circularity, and from 0 μm to 2.0 μm for true position. The average residual values for the neural network model for diameter deviation, circularity, and true position are 1.9 μm, 1.2 μm, and 0.8 μm, respectively. For the linear model, the residual values range from 0.3 μm to 12.2 μm for diameter deviation, from 0 μm to 6.6 μm for circularity, and from 0.2 μm to 2.8 μm for true position. The average residual values for the linear model for diameter deviation, circularity, and true position are 5.5 μm, 2.2 μm, and 1.2 μm, respectively. It can be concluded that the proposed monitoring system based on PCA and MLP networks can provide a very high degree of accuracy in predicting the geometric variability of manufactured parts, given multistage manufacturing data such as material conditions, tempering temperature and tool vibration, and, thus, the volume of dimensional inspections can be minimized.

For the linear models, the coefficient of (multiple) determination R^2 was, on average, 0.746 for diameter deviation, 0.554 for circularity, and 0.673 for true position. The adjusted $\tilde{R}^2$ was, on average, 0.617 for diameter deviation, 0.323 for circularity, and 0.506 for true position. The coefficient of determination R^2 is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{m}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{m}(y_i - \bar{y})^2}, \quad (6)$$

with $0 < R^2 < 1$, where $\bar{y}$ is the mean of the observed data. The corrected/adjusted $\tilde{R}^2$ is given by:

$$\tilde{R}^2 = 1 - \frac{\frac{1}{m-n-1}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\frac{1}{m-1}\sum_{i=1}^{m}(y_i - \bar{y})^2}, \quad (7)$$

with $\tilde{R}^2 < R^2$. However, by adding additional process performance indicators in the proposed methodology such as cutting forces, both the R^2 and the adjusted $\tilde{R}^2$ may increase. Table 5 shows the results for the terms included in the linear model for circularity and fold 1 using a Bayesian parameter estimation approach with a diffuse prior distribution, $p(\boldsymbol{a}, \sigma^2) \propto 1/\sigma^2$. The first column of Table 5 includes the model terms, the second column includes the mean value of the parameter estimates, the third column includes the Standard Error (SE), the fourth column includes the 95% Bayesian equal-tailed Credible Interval (CI), and the fifth column includes the posterior probability that the parameter is greater than zero.

Table 5. Estimated model parameters for circularity and fold 1.

Model terms	Estimate	SE	Bayesian CI	Positive
a_0	0.03815	0.00040	[0.0374, 0.0390]	1.000
a_1	0.00875	0.00229	[0.0042, 0.0133]	0.999
a_2	-0.00680	0.00521	[-0.0172, 0.0036]	0.085
a_3	0.01932	0.00543	[0.0085, 0.0302]	0.998
a_4	-0.00639	0.01267	[-0.0317, 0.0189]	0.288
σ^2	2.1×10^{-6}	1.5×10^{-6}	$[7.1 \times 10^{-7}, 5.7 \times 10^{-6}]$	1.000

A Bayesian formulation treats the model parameters as random variables rather than fixed, unknown quantities and allows one to incorporate prior knowledge about them and thus, leading to the derivation of more flexible models. With the diffuse prior distribution $p(\boldsymbol{a}, \sigma^2) \propto 1/\sigma^2$, the posterior distribution of $\boldsymbol{a}$ conditional on σ^2 is a $(n+1)$-dimensional multivariate normal distribution centered over the vector $\boldsymbol{a}$ of least-squares estimates with covariance matrix $\sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$ and thus, obtaining results identical to that obtained from the frequentist parameter estimation approaches. The marginal

posterior distribution of σ^2 has a scaled inverse-$\mathcal{X}^2$ form, $\sigma^2|\boldsymbol{y}, \boldsymbol{X} \sim \text{Inv} - \mathcal{X}^2\ (m-n-1, \hat{\sigma}^2)$.

5. Conclusions and suggestions for future work

A current trend in the manufacturing industry is the development of flexible and robust product health monitoring systems in order produce accurate products with tight tolerances at the lowest cost and in the lowest processing time. Therefore, non-value-adding operations such as inspection need to be minimized. Although manufacturing processes are usually equipped with multiple sensors and software systems to monitor and control critical process variables such as temperature and vibration, dimensional inspection systems are required to evaluate the end product conformance

This paper has been concerned with the development and application of machine learning algorithms for modelling the geometric variability of manufactured parts, given multistage manufacturing data including material conditions and process monitoring data such as tempering temperature and tool vibration. The proposed methodology can minimize non-value-adding operations such as inspection and allow one to stop the manufacturing process when conformance to a tolerance is not achievable with the current process setting and conditions. Product condition monitoring systems can also be used to initiate the adjustment of certain process parameters such as machining feed rate, spindle speed, and depth of cut thereby preserving optimal manufacturing conditions. If a defective product can be anticipated at an early manufacturing stage and promptly corrected (the manufacturing process), scrap and non-conformance costs can be greatly reduced.

In this research work, PCA was used to reduce the number of variables in the input dataset and avoid overfitting. Although PCA is a well-established method for identifying a low-dimensional structure in high-dimensional datasets, it assumes linear dependencies between the variables. Therefore, future work can consider nonlinear dimensionality reduction methods such as Kernel PCA (KPCA) and self-supervised neural networks to extract the nonlinear structures of the dataset.

Acknowledgements

The authors gratefully acknowledge funding for this research from the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant Reference: EP/P006930/1.

References

[1] Kang HS, Lee JY, Choi S, Kim H, Park JH, Son JY, et al. Smart manufacturing: past research, present findings, and future directions. Int J Prec Eng Man Gr Tech 2016;3:111-128.
[2] Murphy KP. Machine learning: a probabilistic perspective. MIT press; 2012.
[3] Wuest T, Weimer D, Irgens C, Thoben KD. Machine learning in manufacturing: advantages, challenges, and applications. Prod Manuf Res 2016;4:23-45.
[4] McLeay TE. Unsupervised Monitoring of Machining Processes. PhD Thesis. University of Sheffield; 2016.
[5] Zhou X, Belkin M. Semi-supervised learning. Academic Press Library in Signal Processing 2014;1:1239-1269.
[6] Zhong RY, Xu C, Chen C, Huang GQ. Big data analytics for physical internet-based intelligent manufacturing shop floors. Int J Prod Res 2017; 55:2610-2621.
[7] Lee J, Lapira E, Bagheri B, Kao H. Recent advances and trends in predictive manufacturing systems in big data environment. Manuf Lett 2013;1:38-41.
[8] Hocken RJ, Pereira PH. Coordinate measuring machines and systems. 2nd ed. CRC Press: 2012.
[9] Forbes AB, Minh H. Generation of numerical artefacts for geometric form and tolerance assessment. Int J Metr Qual Eng 2012;3:145-150.
[10] Flack D. CMM measurement strategies. NPL; 2014.
[11] Diamantaras KI. Neural networks and principal component analysis. In: Hu YH, Hwang JN, editors. Handbook of Neural Network Signal Processing.CRC Press;2002:209-246.
[12] Cherkassky V, Mulier FM. Learning from data: concepts, theory, and methods: John Wiley & Sons; 2007.
[13] Jolliffe IT. Principal component analysis 2nd ed. Springer; 2002.
[14] Bishop CM. Pattern recognition and machine learning. Springer; 2006.
[15] Haykin S. Neural networks and learning machines. 3rd ed. Pearson; 2009.
[16] Kovač P, Mankova I, Gostimirović M, Sekulić M, Savković B. A review of machining monitoring systems. J Prod Eng 2011;14:1-6.
[17] Risbood K, Dixit U, Sahasrabudhe A. Prediction of surface roughness and dimensional deviation by measuring cutting forces and vibrations in turning porcess. J Mater Proc Tech 2003;132:203-214.
[18] Kovač P, Rodić D, Pucovsky V, Savković B, Gostimirović M. Application of fuzzy logic and regression analysis for modeling surface roughness in face milling. J Intell Manuf 2013;24:755-762.
[19] Plaza EG, López PN, González EB. Efficiency of vibration signal feature extraction for surface finish monitoring in CNC machining. J Manuf Proc 2019;44:145-157.
[20] Arnaiz-González Á, Fernández-Valdivielso A, Bustillo A, de Lacalle LNL. Using artificial neural networks for the prediction of dimensional error on inclined surfaces manufactured by ball-end milling. Int J Adv Manuf Tech 2016;83:847-859.
[21] Gao RX, Tang X, Gordon G, Kazmer DO. Online product quality monitoring through in-process measurement. CIRP Annals 2014;63:493-496.
[22] Papananias M, McLeay TE, Mahfouf M, Kadirkamanathan V. A Bayesian framework to estimate part quality and associated uncertainties in multistage manufacturing. Comput Ind 2019;105:35-47.
[23] Wang G, Ledwoch A, Hasani RM, Grosu R, Brintrup A. A generative neural network model for the quality prediction of work in progress products. Appl Soft Comput 2019;85:105683.
[24] Li L, Ota K, Dong M. Deep learning for smart industry: Efficient manufacture inspection system with fog computing. IEEE Trans Ind Inform 2018;14:4665-4673.
[25] Shi J. Stream of variation modeling and analysis for multistage manufacturing processes. CRC press; 2006.
[26] Papananias M, Fletcher S, Longstaff AP, Forbes AB. Uncertainty evaluation associated with versatile automated gauging influenced by process variations through design of experiments approach. Prec Eng 2017;49:440-455.
[27] Papananias M, McLeay TE, Mahfouf M, Kadirkamanathan V. An Intelligent Metrology Informatics System based on Neural Networks for Multistage Manufacturing Processes. Procedia CIRP 2019;82:444-449.
[28] Golub GH, Van Loan CF. Matrix computations. 4th ed. The Johns Hopkins University Press; 2013.
[29] Forbes AB. Parameter estimation based on least squares methods. In: Pavese F, Forbes AB, editors. Data modeling for metrology and testing in measurement science. Birkhäuser Boston; 2009:147-176.