White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/166957/

Version: Published Version

Article:
Papananias, M., McLeay, T.E., Obajemu, O. et al. (2020) Inspection by exception: a new machine learning-based approach for multistage manufacturing. Applied Soft Computing, 97 (Part A). 106787. ISSN: 1568-4946

https://doi.org/10.1016/j.asoc.2020.106787

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Applied Soft Computing Journal 97 (2020) 106787

Contents lists available at ScienceDirect

Applied Soft Computing Journal

journal homepage: www.elsevier.com/locate/asoc

Inspection by exception: A new machine learning-based approach for multistage manufacturing

Moschos Papananias [a,*], Thomas E. McLeay [b], Olusayo Obajemu [a], 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

ARTICLE INFO

Article history:
Received 5 December 2019
Received in revised form 19 August 2020
Accepted 9 October 2020
Available online 14 October 2020

Keywords:
Artificial Neural Network (ANN)
Fuzzy C-Means (FCM)
Intelligent/smart manufacturing
Machine learning
Multistage Manufacturing Process (MMP)
Principal Component Analysis (PCA)

ABSTRACT

Manufacturing processes usually consist of multiple different stages, each of which is influenced by a multitude of factors. Therefore, variations in product quality at a certain stage are contributed to by the errors generated at the current, as well as preceding, stages. The high cost of each production stage in the manufacture of high-quality products has stimulated a drive towards decreasing the volume of non-added value processes such as inspection. This paper presents a new method for what the authors have referred to as 'inspection by exception' – the principle of actively detecting and then inspecting only the parts that cannot be categorized as healthy or unhealthy with a high degree of certainty. The key idea is that by inspecting only those parts that are in the corridor of uncertainty, the volume of inspections are considerably reduced. This possibility is explored using multistage manufacturing data and both unsupervised and supervised learning algorithms. A case study is presented whereby material conditions and time domain features for force, vibration and tempering temperature are used as input data. Fuzzy C-Means (FCM) clustering is implemented to achieve inspection by exception in an unsupervised manner based on the normalized Euclidean distances between the principal components and cluster centres. Also, deviation vectors for product health are obtained using a comparator system to train neural networks for supervised learning-based inspection by exception. It is shown that the volume of inspections can be reduced by as much as 82% and 93% using the unsupervised and supervised learning approaches, respectively.

© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Metal manufacturing processes usually involve a series of processing stages to achieve the desired geometry and properties of parts. There are various manufacturing methods including casting, forming and machining [1]. Each manufacturing method has its own advantages and disadvantages and the selection of the appropriate technology depends largely on the specific application. Forming and casting processes are mostly followed by machining operations to obtain the final geometry and surface finish of parts. Also, in many manufacturing applications, heat treatment techniques such as quenching, tempering and annealing are employed to modify the physical and mechanical properties of the workpiece. Thus, due to the multistage nature of a typical manufacturing process for metallic parts, the part quality deviations from the nominal geometry at a certain processing stage are contributed by multiple error sources introduced by the current, as well as previous, processing stages [2].

In traditional manufacturing, machining operations are usually followed by dimensional inspection to evaluate part tolerances. There is a variety of dimensional inspection methods including dedicated gauging, On-Machine Probing (OMP) and Coordinate Measuring Machine (CMM) measurement. Dedicated gauging can enable fast feedback to the production loop but it requires an operator to perform the measurements usually with multiple different gauges in order to evaluate all the specified part tolerances and thus, leading to additional high costs for calibrating each hard gauge. OMP refers to the use of Computer Numerically Controlled (CNC) machine tool as a CMM by using a machine tool probe. OMP possesses the advantage of in-situ inspection and thus, allowing machining and inspection with a single workholding setup as well as immediate re-work of the part when required. However, OMP suffers from significant measurement uncertainties due to the large range of complex influence factors and fails to detect machine tool error-induced deviations [3]. Therefore, supplementing with independent measurements, such as CMM measurement, is usually required. CMMs are accurate measurement systems but most require thermally controlled environments to guarantee their measuring capability. In addition,

* Corresponding author.
E-mail address: m.papananias@sheffield.ac.uk (M. Papananias).

https://doi.org/10.1016/j.asoc.2020.106787
1568-4946/© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

the heavyweight structure of these machines gives rise to hysteresis error and results in conservative scanning speeds being selected in order to reduce the dynamic effects [4]. To deal with the machine's dynamic and thermal errors, software technology for error compensation has been proposed and applied to CMMs but software error compensation solutions increase the already high cost of these measurement systems. This research work employs a shop floor Coordinate Measuring System (CMS) based on the parallel kinematic configuration. Parallel Kinematic Machines (PKMs) possess many advantages over Cartesian ones such as better dynamic performance and speed capability but they have a limited operational workspace and nonlinear behaviour across it.

In recent years, manufacturers have faced many challenges to remain competitive with respect to costs, quality, delivery, flexibility, adaptability and sustainability. The trend towards autonomous and intelligent manufacturing systems, relies upon efficient data analytics tools such as those that enable process and product health monitoring and control. The fourth industrial revolution, known as Industry 4.0, concerns the digital transformation of manufacturing processes by integrating manufacturing equipment and systems with data analytics to enable production machines to take decisions based on available data and machine learning algorithms. In particular, modern manufacturing processes are supported by numerous data sources from models and monitored processes. However, the efficient use of these datasets requires statistical techniques such as Principal Component Analysis (PCA) to extract the useful information. PCA is a matrix factorization technique used to reduce the dimensionality of a dataset and reveal hidden informative underlying variables. Over the years, many monitoring systems for machining have been proposed to detect early abnormal process behaviour and reduce product quality variations. Two of the most studied areas of machining process monitoring are the cutting-tool and part condition monitoring based on sensor signals such as force and vibration [5]. Continuous tool monitoring is of high importance especially for difficult-to-machine materials where significant variation in tool life is observed. Most proposed methods are based on supervised learning algorithms though unsupervised learning has also been applied [6]. In particular, machine learning algorithms are largely classified into supervised and unsupervised, based on the mechanism by which the learning process is achieved. Unsupervised learning is a class of machine learning techniques that learn from unlabelled data; only input samples are available. Unsupervised learning techniques can be used in various applications such as dimensionality reduction, clustering and anomaly detection. Supervised learning requires labelled data (input–output samples) for training an algorithm. With the continuously growing amount of data collected from all the production process stages at which each product goes through, there exists a gap in the literature for intelligent product condition monitoring considering the dimensional accuracy of the product and the multistage manufacturing scenario. In addition, while many intelligent condition monitoring systems have been developed for manufacturing processes, most are based on supervised machine learning algorithms. Supervised monitoring systems for fault detection suffer from higher training costs since labelled training data (e.g. post-process inspection results) are required. The present work focuses on reducing the volume of dimensional inspections using multistage manufacturing data and machine learning techniques including both unsupervised and supervised; PCA, Fuzzy C-Means (FCM), and Artificial Neural Networks (ANNs).

Section 2 presents a detailed review of literature relating to monitoring and control methods for manufacturing processes. Section 3 describes the proposed method for inspection by exception. Section 4 presents the basic theory for the algorithms employed in the present work. Section 5 presents the experimental work. Section 6 implements and validates the proposed method for inspection by exception based on both unsupervised and supervised learning. Section 7 presents the concluding remarks.

2. Related literature

Manufacturing enterprises are currently confronted with many challenges as a consequence of growing demands for higher quality of finished products, shorter manufacturing times, greater product complexity and variety, and reduced manufacturing costs, energy consumption and material waste in production. To meet these needs, the refinement of existing monitoring and control systems is becoming increasingly important for intelligent, autonomous manufacturing processes. Manufacturing includes a variety of processes and systems. In CNC machining processes, errors can be broadly classified into two main categories: static or quasi-static and dynamic [2]. Quasi-static errors refer to the static or slow-varying errors and include geometric and kinematic errors, thermal errors, cutting force induced errors, fixturing errors, starting material and tooling inaccuracies, etc. Dynamic errors are much larger and more dependent on the particular process conditions than static errors and are typically caused by sources such as controller error and machine structure vibration. A production process for metallic parts or products usually involves multiple stations or operations such as forming and heat treatment, subtractive machining, in-process and post-process inspection, assembly, and testing. Therefore, in a Multistage Manufacturing Process (MMP), workpiece geometric deviations at a certain production stage are caused by the variation sources introduced by the current stage, as well as the variation propagated from preceding stages [7]. In this paper, we focus on MMPs consisting of heat treatment operations, metal-removing operations, including milling and drilling, conducted on CNC machine tools, and post post-process inspection using automated comparator gauges calibrated through CMM measurements on a master part. Manufacturing operations can be divided into two types: value adding operations and non-value adding operations. For example, machining operations are value adding because they add value to the workpiece by changing its shape, dimensions and surface finish, while inspection operations such as OMP are non-value adding, though they can provide significant advantages to the manufacturing industry in terms of productivity and scrap levels.

With the increasing complexity of manufacturing processes employed to change the geometry and certain properties of a workpiece, conventional CNC approaches may not be able to achieve the desired results in terms of dimension, form and geometry. Dimensional product variation management and reduction for MMPs have been studied extensively and several modelling techniques have been proposed over the years, particularly linearized Stream of Variation (SoV) modelling methods based on differential motion vectors, equivalent fixture error, and kinematic analysis [8]. SoV is a model-based method that utilizes mathematical models such as state-space models to describe the dimensional variation and propagation in multistage assembly processes and multistage machining processes. The derivation of such models is based upon physical knowledge and/or process monitoring data [9]. Loose et al. [10] developed a state-space variation propagation model to describe the product dimensional variation propagation among multiple machining operations with different setups. Their modelling approach can handle general fixture layouts, but they limited the scope of the model only to setup errors. Bazdar et al. [11] focused on diagnosing faults within multistage machining processes using state-space variation propagation modelling and discriminant analysis of setup

errors. Du et al. [12] presented a generic framework for variation propagation modelling for multistage turning processes of rotary workpieces based on differential motion vectors. Wang et al. [13] described a generic variation propagation framework incorporating the elastic deformation variations into state-space modelling for multistage machining processes and Variable Stiffness Structure (VSS) workpieces as most existing SoV methodologies assume that the workpiece is a rigid body. The validation results obtained from a case study concerned with a four-cylinder engine block indicated that the prediction errors are significantly lower than those obtained from conventional SoV modelling methodologies. Although a large amount of research works have showed the reliability of SoV modelling methods for multistage machining processes to predict the dimensional product quality, the applicability of this approach is limited due to the challenges and difficulties associated with constructing and utilizing the SoV model for many MMPs.

In recent years, manufacturing systems have reaped considerable benefit from advances in sensor and information technologies. Therefore, advanced process and product health monitoring and control techniques based on machine learning models and sensor signals such as temperature, force and vibration have attracted a lot of interest. The use of product health monitoring systems can allow us to identify issues associated with the product being manufactured before post-process inspection and thus can greatly help reduce the need for screening inspection without sacrificing the quality of the manufactured product. Wang et al. [14] proposed Multilayer Feedforward Neural Networks (MFNNs), based on an autoencoder for dimensionality reduction, to detect defective products from a powder metallurgy process. An ANN is a collection of interconnected neurons that are able to learn incrementally from their experience to solve complex problems such as nonlinear function approximation. Li et al. [15] proposed a deep learning-based classification model to detect defective products using the concept of fog computing in order to deal with large amounts of data. Papananias et al. [16] presented a Bayesian approach to estimate the results of post-process inspection given in-process inspection data. For turning processes, Salgado et al. [17] proposed an in-process surface roughness prediction system, based on Least Squares Support Vector Machines (LS-SVMs), that uses as inputs feed rate, cutting speed, depth of cut, tool geometry parameters and information extracted from vibrations signals using Singular Spectrum Analysis (SSA). Özel and Karpat [18] used MFNNs to predict both surface roughness and tool flank wear in finish dry hard turning using as inputs material hardness in Rockwell-C scale, cutting speed, feed rate, axial cutting length and the mean values of three force components. For milling processes, Huang [19] presented an intelligent neural-fuzzy in-process surface roughness monitoring system for an end-milling operation using five inputs including spindle speed, feed rate, depth of cut, the average resultant peak force and the absolute average force. Kovac et al. [20] applied fuzzy logic and regression to predict surface roughness in dry face milling using as inputs cutting speed, feed rate, depth of cut and flank wear land width. Han et al. [21] presented a varying-parameter drilling method to improve manufacturing efficiency in successive drilling operations and hole surface quality for multi-hole components. They developed Radial Basis Function (RBF) neural networks to predict surface roughness using spindle speed, feed rate, crater wear, flank wear, outer corner wear, thrust force and torque.

Published research on dimensional product health monitoring is limited and much of it focuses on monitoring only the machining process to identify the end product quality, though manufacturing processes typically involve multiple production stages. In addition, although the subject of manufacturing process monitoring and control is a well-developed field of intelligent manufacturing, the manufacturing industry has adopted few monitoring and control systems to replace decision making of a human with a machine. The robustness issues when operating under different conditions and the high costs of training supervised monitoring systems remain two of the major issues faced when extending academic research solutions to industrial exploitation. There is therefore a need for intelligent monitoring and control systems that are able to function under various process conditions and conditions of uncertainty. This paper presents a new method, referred to as 'inspection by exception'. The proposed method is based on the idea of predicting the end product quality using machine learning and multistage, in-process monitoring data in order to capture sufficient knowledge about the production process and then inspecting the product only if it cannot be classified as conforming or non-conforming with a high degree of certainty. The proposed methodology is tested on a MMP consisting of different processing stages including heat treatment and machining, and on different dimensional metrology characteristics including diameter, true position, and circularity. This modelling problem can be considered to be representative of many manufacturing processes, particularly small batch manufacturing sectors such as aerospace manufacturing applications due to the nonlinearity, high dimensionality, sparsity and uncertainty of the manufacturing process and workpiece data. In addition, in order to achieve inspection by exception with the minimal cost of implementation, clustering-based inspection by exception is also proposed and validated.

3. Inspection by exception method

This section describes the proposed method for inspection by exception. The key idea of the method is that inspection is not required for parts that can be categorized as healthy or unhealthy with a high degree of certainty but only for those parts that are in the corridor of uncertainty so that the volume of inspections can be reduced without making a wrong decision, e.g. rejecting a part that conforms to design specifications (Type I error) or accepting a part that does not conform to design specifications (Type II error).

The part quality characteristics considered in this paper are the diameter deviation, true position and circularity of a bore. Table 1 shows the deviations from nominal values for the three quality characteristics of interest obtained from the Equator gauge in scanning mode using the CMM Compare method. The columns of Table 1 are as follows: the first column includes the part label/number; the second column includes the diameter deviation; the third column includes the true position; the fourth column includes the circularity; and the fifth column indicates which parts conform to the specified tolerances. The deviations from (drawing) nominal values have been obtained by calculating the absolute difference between the actual/measured value and the nominal value for each measurand and thus, $h_j = \left|y_j - \tilde{y}_j\right|$, for $j = 1, 2, \ldots, n$, where y_j denotes an observation of the measured quantity Y_j and $\tilde{y}_j$ denotes the (drawing) nominal value of the same quantity, Y_j. Suppose the tolerances are ± 0.0700 mm for diameter, 0.0100 mm for true position (47.5 $\pm$ 0.075 in X axis from datum B and 40 $\pm$ 0.075 in Y axis from datum A), and 0.0500 mm for circularity. The parts that do not conform to the specified tolerances are the parts labelled as: 1, 10, and 23. The diameter deviations for these three parts are: 0.0771 mm, 0.0705 mm, and 0.0802 mm, respectively. The fourth largest diameter deviation is 0.0623 mm for part 19.

The decision rule of the proposed method is given by:

$$If\ \hat{h}_{jp} \begin{cases} < \underline{\tau}_j & Healthy\ parts - No\ inspection \\ Otherwise & Inspect\ parts \\ > \overline{\tau}_j & Unhealthy\ parts - No\ inspection \end{cases} \quad (1)$$

Table 1
Product quality deviations obtained from the Equator gauge.

Parts	Diameter deviation (mm)	True position (mm)	Circularity (mm)	Conforming parts
1	**0.0771**	0.0075	0.0429	No
2	0.0573	0.0036	0.0364	Yes
6	0.0604	0.0032	0.0370	Yes
7	0.0550	0.0045	0.0355	Yes
9	0.0502	0.0039	0.0354	Yes
10	**0.0705**	0.0061	0.0420	No
12	0.0509	0.0016	0.0379	Yes
13	0.0574	0.0046	0.0406	Yes
14	0.0579	0.0019	0.0389	Yes
15	0.0564	0.0029	0.0365	Yes
16	0.0500	0.0060	0.0368	Yes
18	0.0525	0.0046	0.0363	Yes
19	0.0623	0.0061	0.0398	Yes
21	0.0598	0.0029	0.0364	Yes
22	0.0509	0.0042	0.0367	Yes
23	**0.0802**	0.0055	0.0411	No
24	0.0542	0.0028	0.0410	Yes

Two implementations are presented in this paper, both using machine learning but other modelling techniques may be applicable and may also be possibly advantageous to implement the method. In particular, a neural network-based approach is developed to predict the product health metric deviation vector, $\hat{\boldsymbol{h}}_j = \left(\hat{h}_{j1}, \hat{h}_{j2}, \ldots, \hat{h}_{j\tilde{P}}\right)^{\mathrm{T}}$, for $p = 1, 2, \ldots, \tilde{P}$, and assess the requirement for inspection, where $\tilde{P}$ denotes the number of parts used to test the model. The proposed method allows users freedom to choose their preferred bounds $\underline{\tau}_j$ and $\overline{\tau}_j$. However, if the bounds $\underline{\tau}_j$ and $\overline{\tau}_j$ are not selected appropriately, then the gap between $\underline{\tau}_j$ and $\overline{\tau}_j$ may be too large, resulting in unnecessary inspections or the gap may be too small, which may result in healthy parts being rejected (Type I error) and unhealthy parts being accepted (Type II error). The gap between the bounds depends on the accuracy of the model and the uncertainty associated with the directly measured product health metric deviation vector, h_{jp}.

Also, a clustering-based approach is developed to achieve inspection by exception based on the normalized Euclidean distances between the principal components and cluster centres. This approach is based on the FCM algorithm, which allows each data point to belong to more than one cluster. The unsupervised approach will be presented first in this work.

4. Basic theory

4.1. Principal component analysis

The main purpose of PCA is to extract the useful information from a dataset consisting of a set of correlated variables and to represent this information as a new set of uncorrelated variables [22,23]. These uncorrelated variables are principal components and are the directions in which the data have the largest variances. Suppose that we have observations on N variables $x_1, x_2, \ldots, x_N$. The first principal component, y_1, is defined to be the linear combination $a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N = \sum_{j=1}^{N} a_{1j}x_j$, $\|\boldsymbol{a}_1\| = 1$, with maximum variance. The second principal component, y_2, is a linear combination $a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N = \sum_{j=1}^{N} a_{2j}x_j$, $\|\boldsymbol{a}_2\| = 1$, which is uncorrelated with the first derived variable and has maximum variance. The Kth principal component, y_K, for $K = 1, 2, \ldots, N$, can be defined as a linear combination $a_{K1}x_1 + a_{K2}x_2 + \cdots + a_{KN}x_N = \sum_{j=1}^{N} a_{Kj}x_j$, $\|\boldsymbol{a}_K\| = 1$, which is uncorrelated with the first $K-1$ derived variables and which has maximum variance. Let $\boldsymbol{\Sigma}$ be the covariance matrix of the vector of variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)^{\mathrm{T}}$, which we assume to be positive definite. Then, the first principal component can be obtained by finding the weight vector $\boldsymbol{a}_1$ of N constants that maximizes the variance of $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{x}$, given that $\|\boldsymbol{a}_1\| = 1$. The technique of Lagrange multipliers can be used to maximize $V(y_1) = \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}_1$ subject to $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1 = 1$. The second principal component can be obtained by finding the weight vector $\boldsymbol{a}_2$ of N constants that maximizes the variance of $\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{x}$, given that $\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{a}_2 = 1$ and $\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}_1 = 0$. Hence, the Kth principal component is given by $y_K = \boldsymbol{a}_K^{\mathrm{T}}\boldsymbol{x}$ where $\boldsymbol{a}_K$ is an eigenvector of $\boldsymbol{\Sigma}$ corresponding to its Kth largest eigenvalue $\lambda_K = V(y_K)$. The proportion of the total variability explained by the first K principal components can be calculated as:

$$\frac{\sum_{j=1}^{K} \lambda_j}{\sum_{j=1}^{N} \lambda_j}. \tag{2}$$

4.2. Fuzzy C-means clustering

Clustering algorithms find natural groupings in data. Clustering methods utilize distance functions to measure the similarity of data points. The FCM algorithm is a soft clustering method in which each data point can belong to multiple clusters with varying degrees of membership [24]. The distance between data points and cluster centres can be computed by the Euclidean distance:

$$d_{ij} = \left\| x_i - c_j \right\|, \tag{3}$$

where $x_i \in \Re^N$, for $i = 1, 2, \ldots, M$, is the ith data point of Mobservations in N-dimensional Euclidean space and $c_j \in \Re^N$, for $j = 1, 2, \ldots, C$, is the centre of the jth cluster. FCM is based on the minimization of the following objective function:

$$J_w = \sum_{i=1}^{M} \sum_{j=1}^{C} \mu_{ij}^{w} d_{ij}^2, \qquad 1 < w < \infty, \tag{4}$$

subject to the constraints $\mu_{ij} \in [0, 1]$ and $\forall i$: $\sum_{j=1}^{C} \mu_{ij} = 1$, where μ_{ij} is the degree of membership of x_i in the jth cluster and w is the weighting exponent which controls the degree of fuzziness. Note that the cluster centres are calculated by:

$$c_j = \frac{\sum_{i=1}^{M} \mu_{ij}^{w} x_i}{\sum_{i=1}^{M} \mu_{ij}^{w}}. \tag{5}$$

The FCM algorithm in MATLAB performs the following steps during clustering:

i The cluster membership values, μ_{ij}, are randomly initialized.
ii The cluster centres, c_j, are calculated by Eq. (5):
iii The cluster membership values, μ_{ij}, are updated according to:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{w-1}}}$$

iv The objective function, J_w, is calculated by Eq. (4).

Steps ii–iv are repeated until J_w improves by less than a predefined threshold value between two consecutive iterations or until a predefined number of iterations has been reached.

4.3. Artificial neural networks

ANNs are computational models that emulate the learning process of biological neural networks. Over the years, many different types of ANNs have been proposed, one of the most popular being the MFNNs also known as Multi-Layer Perceptron (MLP) networks. Let L be the number of layers of the network. Each layer

Table 2
Chemical composition of starting material.

Carbon C (%)	Silicon Si (%)	Manganese Mn (%)	Phosphorus P (%)	Sulphur S (%)	Chromium Cr (%)	Nickel Ni (%)	Molybdenum Mo (%)	Copper Cu (%)
0.37	0.23	0.55	0.007	0.001	1.34	1.53	0.22	0.11

is denoted by l, $l = 0, 1, \ldots, L$, with $l = 0$ denoting the input layer, $l = L$ the output layer, and $l = 1, \ldots, L-1$ the hidden layers. The neurons in any layer $l = 1, \ldots, L$ operate as follows:

$$net_j(l) = \sum_{i=1}^{\mathcal{N}(l-1)} w_{ij}(l-1, l)\, O_i(l-1) + \theta_j(l), \quad O_j(l) = f\left(net_j(l)\right), \tag{6}$$

where $O_j(l)$ is the output of neuron j in layer l, f is an activation function, $\mathcal{N}(l-1)$ is the number of nodes in layer $l-1$, $w_{ij}(l-1, l)$ is the synaptic weight coefficient associated with the connection from node i in layer $l-1$ to node j in layer l, $O_i(l-1)$ is the output of node i in layer $l-1$, and $\theta_j(l)$ is the bias of neuron j in layer l. The dataset required to train the network consists of $\overline{P}$ input-desired output patterns $\left\{\boldsymbol{x}^{(1)}, \boldsymbol{h}^{(1)}\right\}, \left\{\boldsymbol{x}^{(2)}, \boldsymbol{h}^{(2)}\right\}, \ldots, \left\{\boldsymbol{x}^{(\overline{P})}, \boldsymbol{h}^{(\overline{P})}\right\}$, where $\boldsymbol{x}^{(p)} = \left[x_1^{(p)}, \ldots, x_{\mathcal{N}(0)}^{(p)}\right]^{\mathrm{T}}$ is the input vector for the pth pattern and $\boldsymbol{h}^{(p)} = \left[h_1^{(p)}, \ldots, h_{\mathcal{N}(L)}^{(p)}\right]^{\mathrm{T}}$ is the desired output vector for the pth pattern. Let $\hat{\boldsymbol{h}}^{(p)} = \left[\hat{h}_1^{(p)}, \ldots, \hat{h}_{\mathcal{N}(L)}^{(p)}\right]^{\mathrm{T}}$ denote the output vector for the pth pattern. The Mean Squared Error (MSE) is given by:

$$J_w = \frac{1}{\overline{P}} \sum_{p=1}^{\overline{P}} e(p), \tag{7}$$

where $e(p) = \left\|\boldsymbol{h}^{(p)} - \hat{\boldsymbol{h}}^{(p)}\right\|^2 = \sum_{j=1}^{\mathcal{N}(L)} \left[h_j^{(p)} - \hat{h}_j^{(p)}\right]^2$. Minimization of J_w is attempted by using a particular Back-Propagation (BP) algorithm to adjust the network's weights [25].

5. Experimental work

Experimental work was performed, involving multiple stages of manufacturing, to produce steel bearing housing parts. The MMP consists of heat treatment, grinding, hardness testing, machining and post-process inspection on the shop floor (see Fig. 1). The drawing and Computer-Aided Design (CAD) model of the part is shown in Fig. 2. The part has several critical features that should be sensitive to manufacturing process conditions and errors. For example, any tool/spindle runout can be indicated in the diameter measurement and variation in material hardness can cause differing levels of deflection. The starting material was EN24T steel. The chemical composition of this material is shown is Table 2. Table 3 shows the mechanical properties of the starting material. The EN24T material, is readily machinable because it is heat treated (hardened and tempered and stress-free annealed) by the supplier. However, the starting material was further hardened and tempered at different temperatures. A VECSTAR furnace was used for heat treatment. In particular, the starting material bar was sawn and machined to a nominal size, 25 off blocks. The twenty-five material blocks were divided into five batches and heated up to 845 °C, separately (five batches of five blocks each), and then quenched in oil for hardening. After the hardening operation, tempering was performed at different temperatures (450 °C, 550 °C and 650 °C) to add variability in the properties of the material such as surface hardness (see Table 4). Five K-type thermocouples with protection sheath were used to measure variation in temperature gradient within the furnace during heat treatment. Following the heat treatment process, the blocks were grinded to improve the quality of material surface and measure the material surface hardness. A Rockwell device was used to measure the surface hardness of the heat treated blocks.

Table 3
Mechanical properties of starting material.

Size (mm)	Rp02 N/mm^2 (MPa)	Rm N/mm^2 (MPa)	A5 (%)	Z (%)	Hardness (HB)
10	766	941	18.5	63	272-296

A full factorial design (see Table 5) was performed for machining using a DMG MORI NVX 5080 3-axis machine. The factors included material surface hardness, feed rate, spindle speed and datum error in both X and Y axes when handing over to the second orientation of the workpiece (flipped around the Y axis) for machining the features of its bottom side. All the factors included two levels and one centre point. Each workpiece was machined with coolant. Six cutting tools were employed to manufacture the steel bearing housing parts and changed when reached a certain flank wear width. The wear was evaluated on each flute using a Leica microscope after obtaining each product. A Kistler quartz multicomponent dynamometer (9255B), consisting of four 3-component force sensors, and DynoWare software were used to measure force. The dynamometer was located between the vice holding the workpiece and the machine table. An accelerometer sensor, placed on the spindle, and NI LabVIEW SignalExpress software were used to obtain vibration data. The sampling rate for both force and vibration data was 10 kHz. In total, eighteen parts were machined (seventeen parts for the experimental design and one part to be used as a master part in CMM Compare measurement).

A Renishaw Equator gauge (300 Extended Height) equipped with the SP25 3-axis analogue scanning probe was used for post-process inspection under workshop conditions [26,27]. The Equator was employed in CMM Compare and in scanning measuring mode. The CMM Compare method requires a production part to be labelled as a 'master' part and measured on a calibrated CMM to produce a calibration file for the comparator measurement system [28]. A Mitutoyo CMM located in a temperature controlled room was used to produce the calibration file. The master part had been thermally stabilized before generating the calibration file. The CMM was equipped with a Renishaw REVO RSP3 3D scanning probe. The stylus used for both the Equator and the CMM was a typical 30 mm long stylus with tungsten carbide stem and a 2 mm diameter ruby ball. The same part fixturing setup was also used for both the Equator and the CMM.

6. Implementation and validation of the proposed method

6.1. Clustering-based inspection by exception method

This section presents an unsupervised learning approach to achieve inspection by exception. The machined parts were clustered into two groups using the FCM clustering algorithm in MATLAB. The dataset includes: the Root-Mean-Square (RMS), sample kurtosis, sample skewness, sample variance and mean features of three vibration components (Vx, Vy, Vz); the same five features (RMS, sample kurtosis, sample skewness, sample variance and

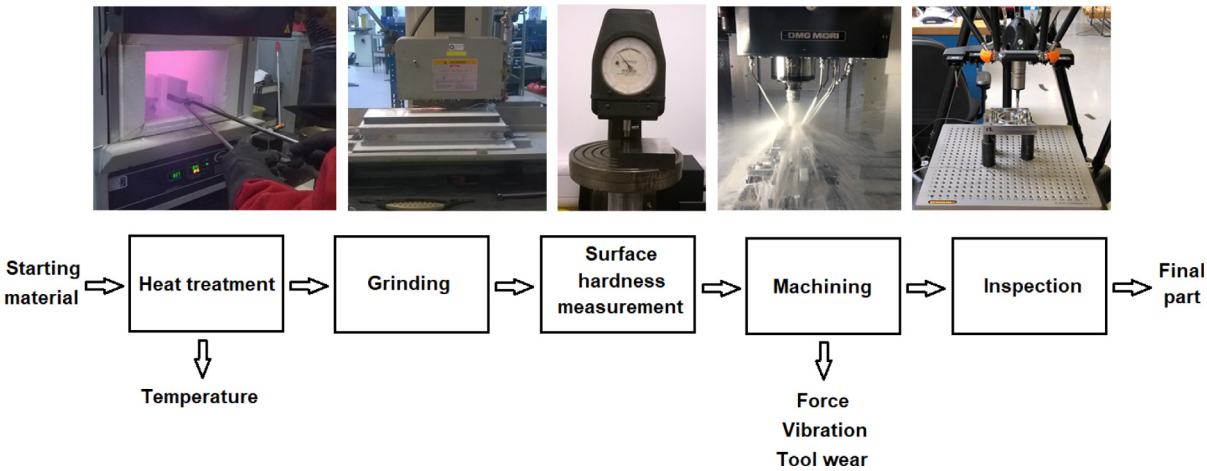

Fig. 1. Multistage manufacturing process.

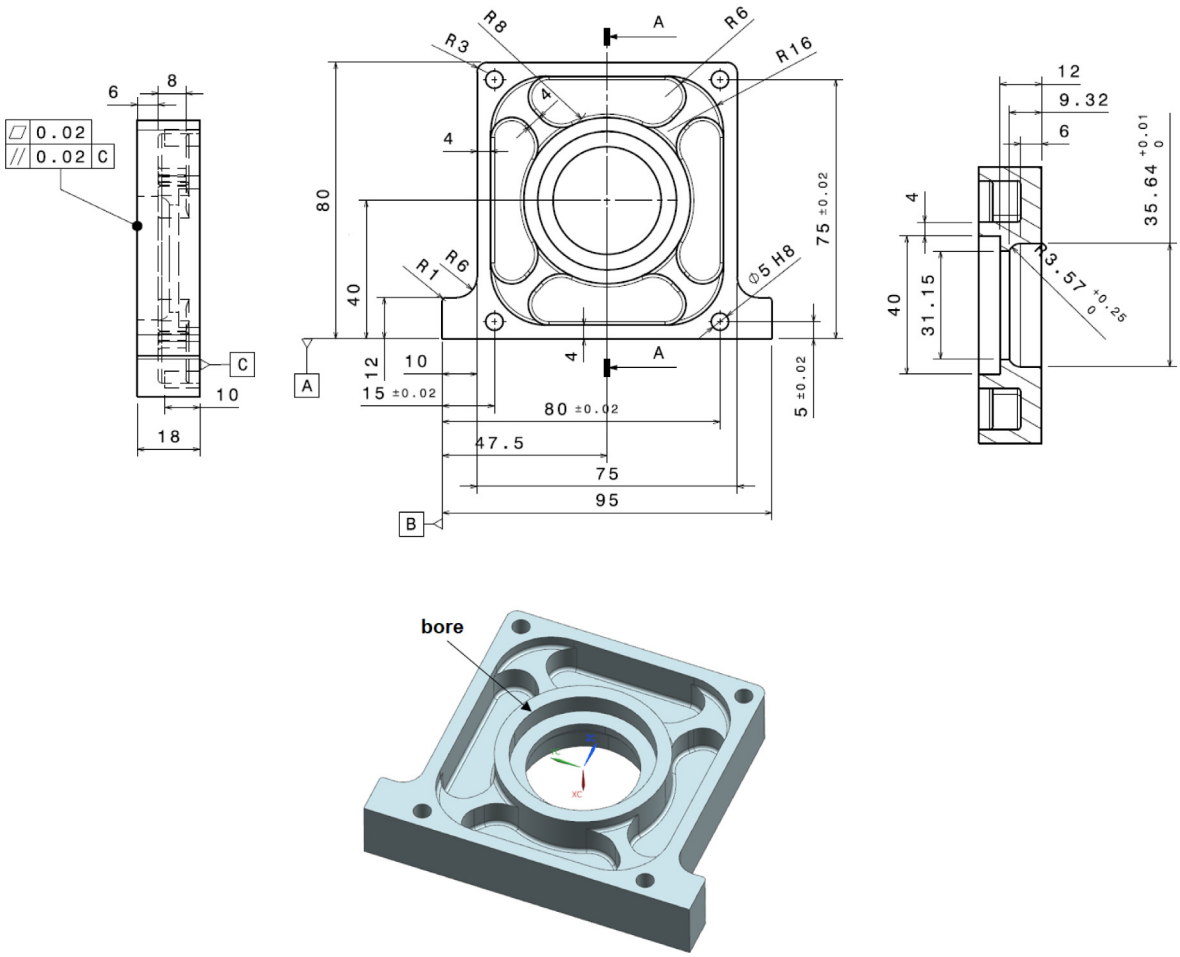

Fig. 2. Drawing and CAD model of the part.

sample mean) of average values of three force components (Fx, Fy, Fz) obtained from all the four sensors of the dynamometer; the maximum temperature obtained from the five K-type thermocouples during tempering; and a coded vector corresponding to the material surface hardness. All the extracted features were normalized by the Euclidean norm (2-norm).

One major difficulty in multivariate analysis is the problem of visualizing high-dimensional data. To reduce the number of variables to a few ($\ll N$) variables that represent most of the information in the original variables, PCA was performed in MATLAB using a Singular Value Decomposition (SVD) of the input data matrix $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}} \in \mathcal{R}^{M\times N}$, where both $\boldsymbol{U} \in \mathcal{R}^{M\times M}$ and $\mathbf{V} \in \mathcal{R}^{N\times N}$ are orthogonal matrices and $\boldsymbol{\Sigma} \in \mathcal{R}^{M\times N}$ is a diagonal matrix with diagonal elements, σ_j, such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(M,N)} \geq 0$. Therefore, mean-centring the columns of the normalized input data matrix $\boldsymbol{X}$ was an essential pre-processing step in the process of dimensionality reduction. A 2-fold cross-validation approach was employed to test the PCA-based clustering approach using

Table 4
Heat treatment.

	Block number	Hardening (°C)	Tempering (°C)
1	13	845	450
2	14	845	550
3	10	845	450
4	12	845	650
5	22	845	450
6	15	845	650
7	9	845	650
8	19	845	450
9	16	845	450
10	24	845	650
11	21	845	650
12	7	845	450
13	2	835	650
14	18	845	650
15	6	845	650
16	23	845	450
17	1	835	450

Table 5
Full factorial design.

Block number	Surface hardness	Feed rate	Spindle speed	Datum error in both X and Y (mm)
13	Hard	Programmed	+20%	0
14	Middle	+10%	+10%	0.01
10	Hard	+20%	Programmed	0.02
12	Soft	Programmed	Programmed	0.02
22	Hard	Programmed	Programmed	0.02
15	Soft	+20%	+20%	0.02
9	Soft	Programmed	+20%	0.02
19	Hard	Programmed	+20%	0.02
16	Hard	Programmed	Programmed	0
24	Soft	Programmed	Programmed	0
21	Soft	+20%	Programmed	0.02
7	Hard	+20%	+20%	0.02
2	Soft	+20%	+20%	0
18	Soft	Programmed	+20%	0
6	Soft	+20%	Programmed	0
23	Hard	+20%	+20%	0
1	Hard	+20%	Programmed	0

data from the whole experimental design (seventeen manufactured parts). In particular, the dataset of seventeen parts was partitioned into two sub-datasets: one sub-dataset was used for training the clustering algorithm and one sub-dataset was used for testing it. The cross-validation process was repeated two times so that both sub-datasets were used as the validation dataset once. Sub-dataset 1 includes the parts: 1, 2, 6, 7, 9, 10, 12, 13, 14. Sub-dataset 2 includes the parts: 15, 16, 18, 19, 21, 22, 23, 24. The percent variability explained by the first two components are: 81.23% for sub-dataset 1 and 89.61% for sub-dataset 2. For sub-dataset 2 used as test dataset, the mean squared reconstruction error for the training data considering the first two principal components was 0.00071 and for the test data considering the first two principal components obtained from the trained PCA model was 0.00430. Similarly, for sub-dataset 1 used as test dataset, they were 0.00075 and 0.00180, respectively.

Given the first two principal components, the FCM algorithm was employed to partition the data into two clusters. The FCM algorithm initially generates a random membership matrix. In each clustering iteration, the FCM algorithm calculates the cluster centres and updates the membership matrix using the calculated cluster centre locations. The algorithm then computes the objective function value. The clustering process stops when the objective function improvement falls below a predefined threshold value or when the maximum number of iterations has been reached. The amount of fuzzy overlap during clustering was set to 15, the maximum number of iterations was set to 100, and

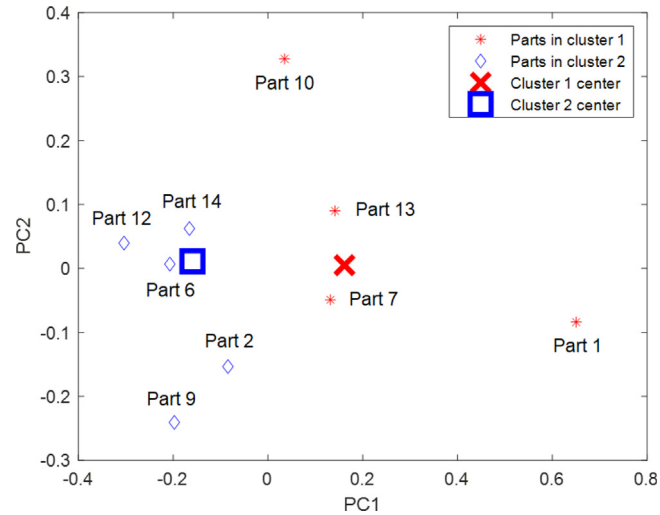

Fig. 3. PCA-based FCM clustering results for fold 1 (training).

the minimum improvement in objective function between two consecutive iterations was set to 0.00001. The clustering algorithm was trained using the first two principal components of the training dataset. Then, new data/test dataset were assigned to the existing clusters and the normalized Euclidean distances between the first two principal components and cluster centres were computed. The computed distances are used to assess whether or not a product meets its specifications and identify the parts that require inspection for conformance assessment. Figs. 3–6 show the clustering results. Figs. 3 and 5 show the clustering results obtained from training the clustering algorithm for fold 1 and 2, respectively, using the first two principal components. Fold 1 uses sub-dataset 1 for training and sub-dataset 2 for testing. Fold 2 uses sub-dataset 2 for training and sub-dataset 1 for testing. Figs. 4 and 6 show the assignment of the test dataset (the first two principal components obtained from the trained PCA model) to the existing clusters shown in Figs. 3 and 5, respectively. Figs. 7 and 8 show the normalized Euclidean distances between the first two principal components and cluster centres for fold 1 and 2, respectively. Based on the results shown in Figs. 7 and 8, it can be concluded that the PCA-based FCM clustering approach can reduce the volume of inspections from seventeen parts to one part for fold 1 and to three parts for fold 2, since only part 19 for fold 1 and parts 10, 13 and 7 for fold 2 require inspection to determine successfully whether or not they conform to specifications.

6.2. Neural network-based inspection by exception method

This section implements a supervised learning approach using PCA-based MLP networks to provide deviation vectors for product health and achieve inspection by exception.

6.2.1. Product health metric vectors

A CMS operating in comparator mode (Equator) was used to obtain m measurements on a quantity Y, regarded as a random variable and called the measurand, e.g. the diameter of a bore. Let $y_1, y_2, \ldots, y_m$ denote the corresponding measured values obtained from the comparator system in reproducibility conditions. The measurements $y_1, y_2, \ldots, y_m$ involve comparator measurements and thus, random effects are dominant [29]. Given the measurand Y and its attached Probability Density Function (PDF) $g(y)$, with $g(y) \geq 0$ for all values of y and $\int_{-\infty}^{\infty} g(y)\, dy = 1$, the

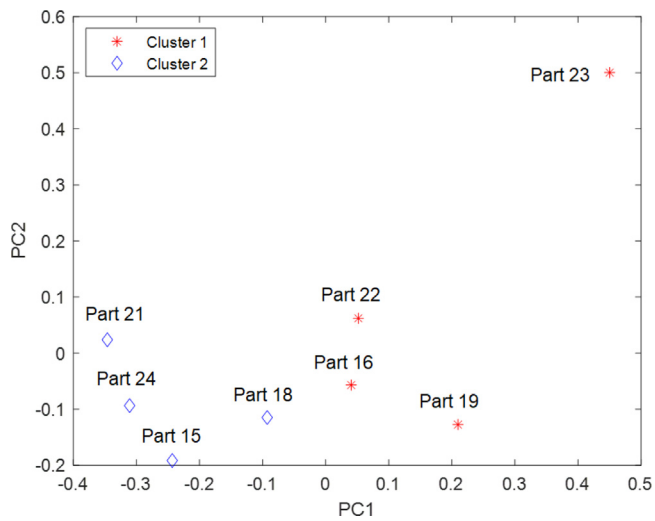

Fig. 4. PCA-based FCM clustering results for fold 1 (testing).

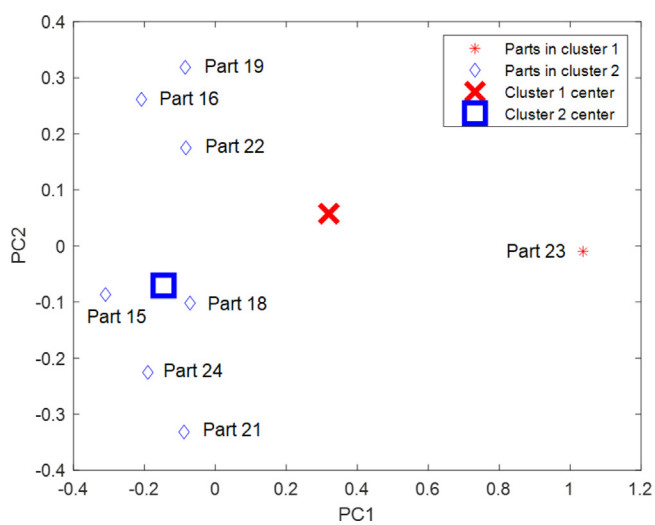

Fig. 5. PCA-based FCM clustering results for fold 2 (training).

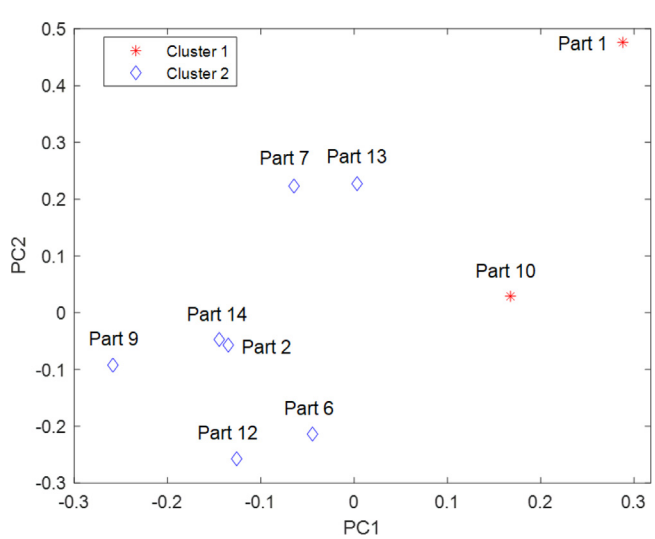

Fig. 6. PCA-based FCM clustering results for fold 2 (testing).

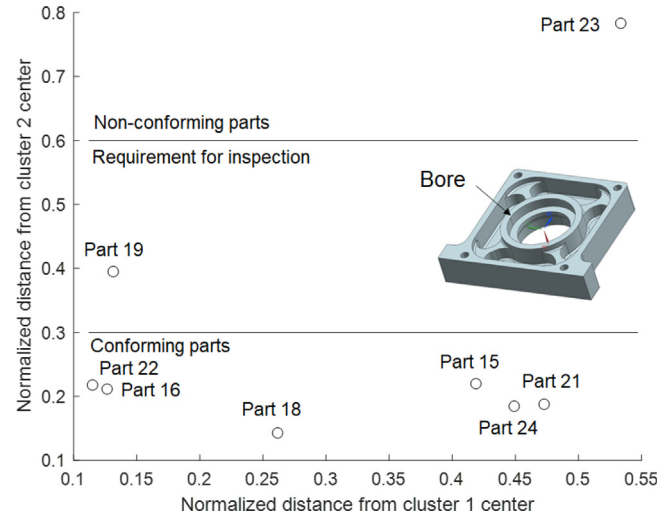

Fig. 7. Normalized Euclidean distances between the first two principal components and cluster centres for fold 1.

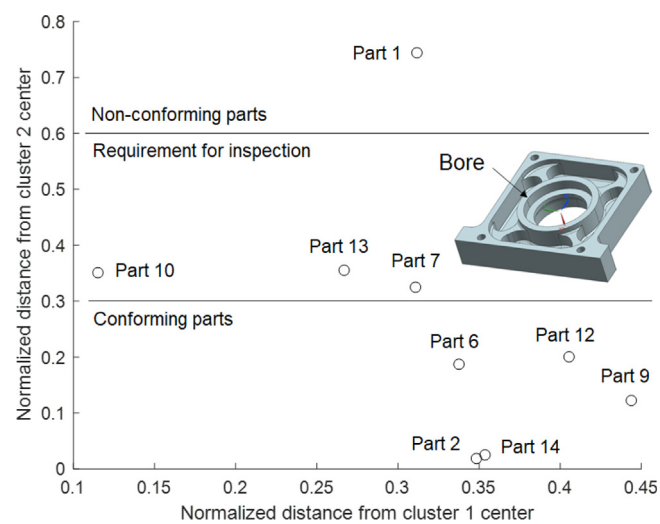

Fig. 8. Normalized Euclidean distances between the first two principal components and cluster centres for fold 2.

expectation $E(Y)$ and the variance $V(Y)$ are defined, respectively, as:

$$E(Y) = \int_{-\infty}^{\infty} y g(y)\, dy, \tag{8}$$

$$V(Y) = E\left[(Y - E(Y))^2\right] = \int_{-\infty}^{\infty} [y - E(Y)]^2 g(y)\, dy, \tag{9}$$

Note that if $Y \sim \mathrm{N}\left(\mu, \sigma^2\right)$ then $E(Y) = \mu$ and $V(Y) = \sigma^2$, where μ is the expectation or mean and σ is the standard deviation, the positive square root of the variance σ^2. For n geometric tolerances applied to one or more features, let $\mathcal{Y}$ denote the product health matrix:

$$\mathcal{Y} = \left(y_{ij}\right) = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \in \Re^{m \times n}, \tag{10}$$

where y_{ij} denotes the element located in the ith row and the jth column of the product health matrix $\boldsymbol{y}$. Given that the product is measured m times independently, under repeatability conditions, estimates for $\mu_1, \mu_2, \ldots, \mu_n$ and $\sigma_1, \sigma_2, \ldots, \sigma_n$ can be obtained by the sample means $\overline{y}_1, \overline{y}_2, \ldots, \overline{y}_n$ and the sample standard deviations $s_1, s_2, \ldots, s_n$, respectively:

$$\overline{y}_j = \frac{1}{m}\sum_{i=1}^{m} y_{ij},\ \ s_j = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}\left(y_{ij}-\overline{y}_j\right)^2},\ \ j = 1, 2, \ldots, n, \tag{11}$$

Note that while the sample variance s^2 is an unbiased estimator of the population variance σ^2, the sample standard deviation s is a biased estimator of the population standard deviation σ [30]. A measure of linear association between the observations of jth and lth product quality characteristics can be provided by the sample covariance:

$$s_{jl} = \frac{1}{m-1}\sum_{i=1}^{m}\left(y_{ij}-\overline{y}_j\right)\left(y_{il}-\overline{y}_l\right), \qquad j = 1, 2, \ldots, n,$$
$$l = 1, 2, \ldots, n. \tag{12}$$

The sample covariance of the standardized observations can be given by the sample correlation coefficient:

$$r_{jl} = \frac{s_{jl}}{\sqrt{s_{jj}}\sqrt{s_{ll}}} = \frac{\sum_{i=1}^{m}\left(y_{ij}-\overline{y}_j\right)\left(y_{il}-\overline{y}_l\right)}{\sqrt{\sum_{i=1}^{m}\left(y_{ij}-\overline{y}_j\right)^2}\sqrt{\sum_{i=1}^{m}\left(y_{il}-\overline{y}_l\right)^2}},$$
$$-1 \le r_{jl} \le 1. \tag{13}$$

Two random variables Y_j and Y_l are statistically independent if their joint PDF $g_{jl}\left(y_j, y_l\right)$ can be factored as $g_{jl}\left(y_j, y_l\right) = g_j\left(y_j\right) g_l\left(y_l\right)$. Independent random variables are also uncorrelated but the converse does not necessarily hold [31]. The product correlation matrix given by:

$$\boldsymbol{R} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1n} \\ r_{21} & 1 & \ldots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \ldots & 1 \end{bmatrix}, \tag{14}$$

can be computed to exploit the correlation between the same characteristics for identical features.

6.2.2. Experimental evaluation of uncertainty associated with equator CMM compare measurement

This section presents an experimental method to evaluate the measurement uncertainty associated with an estimate of a product quality characteristic in order to provide confidence in the comparator measurement results and allow the calculation of conformance and non-conformance probabilities. The measurement uncertainties associated with Equator CMM Compare method were calculated experimentally considering: (i) the standard uncertainty, $u(cal)$, associated with the uncertainty of the calibration of the master part; ii) the standard uncertainty, $u(p)$, associated with the comparative coordinate measurement procedure; and iii) the standard uncertainty, $u(b)$, associated with the systematic error component, b, of the comparative coordinate measurement [32]. The effect of the uncertainty in the Coefficient of Thermal Expansion (CTE) will not be considered in this evaluation procedure. The combined standard uncertainty, $u_c\left(y\right)$, of any measurand was calculated as follows:

$$u_c\left(y\right) = \sqrt{u^2(cal) + u^2(p) + u^2(b)} + b. \tag{15}$$

The standard uncertainty, $u(cal)$, was estimated by the experimental standard deviation of the mean, $s\left(\overline{y}_{cal}\right) = s/\sqrt{m_{cal}}$, using the CMM measurements on the master part. The uncertainty components, $u(p)$ and $u(b)$, were calculated by the experimental standard deviation of the mean using the comparative coordinate measurements on the test part and master part, respectively. The systematic error component, b, was calculated by $|\overline{y} - y_{cal}|$, where $\overline{y}$ denotes the mean value of the Equator measurements on the master part and y_{cal} is the CMM calibrated (mean) value of the same part and measurand. The expanded measurement uncertainties, U, of the master part, used to validate the CMM Compare method, are: 0.54 μm, 0.76 μm and 1.08 μm for diameter, true position and circularity, respectively, for a coverage factor $k = 2$ and a confidence level of 95.45%. For example, for the test part 24, the expanded measurement uncertainties are similarly: 0.63 μm, 0.91 μm and 2.66 μm, respectively. If a random sample of size m is drawn from a population which forms a normal distribution, then, the appropriate test statistic is $t = \left(\overline{y} - \mu\right)/\left(s/\sqrt{m}\right)$, which has a student's t-distribution with $\nu = m-1$ degrees of freedom. The mean of the t-distribution is zero and its variance is $\nu/(\nu-2)$ for $\nu > 2$. As $\nu \to \infty$, the t-distribution approaches a normal distribution with $\mu = 0$ and $\sigma = 1$ [30].

6.2.3. Prediction of product health metric deviation vectors

To predict the product health metric deviation vectors, an MLP network with 3 inputs, one hidden layer consisting of five neurons, and one output was developed for each product quality characteristic in MATLAB. For all the models, tan-sigmoid transfer functions were used in the hidden layer and a linear activation function was used for the output neuron. The inputs to the network are the first three principal components of the normalized dataset used for the clustering approach presented in Section 6.1. The output is the product health metric deviation vector for each quality characteristic (diameter, true position and circularity of the bore), obtained by the absolute difference between the (drawing) nominal value and the measured value obtained from the Equator gauge in scanning mode using the CMM Compare method. All the models were trained by Resilient BP, which converges faster than the traditional BP algorithm [33].

A 4-fold cross-validation approach was employed to test the neural network models using data from the manufacture of sixteen parts (block 14 was excluded). In particular, the dataset of sixteen parts was randomly partitioned into four sub-datasets and three sub-datasets were used to train the models and a single sub-dataset was used to test the models. The cross-validation process was repeated four times so that each of the sub-datasets was used as the validation dataset once. The performance of each model was evaluated using the MSE. The training process was stopped when reaching the specified maximum number of epochs (1000) or the validation error began to rise. Sub-dataset 1 includes the parts: 2, 23, 22, 21. Sub-dataset 2 includes the parts: 12, 13, 24, 7. Sub-dataset 3 includes the parts: 19, 6, 9, 10. Sub-dataset 4 includes the parts: 1, 16, 15, 18. Each neural network model was trained using the first three components obtained from the PCA transformation of the training dataset. The percent variability explained by the first three components are: 89.17% for sub-datasets 2,3,4; 90.76% for sub-datasets 1,3,4; 90.88% for sub-datasets 1,2,4; and 93.27% for sub-datasets 1,2,3. Each trained neural network model was tested using the first three components obtained from the trained PCA model on the test dataset. The mean squared reconstruction error for the training dataset using the first three components was 0.00044 for sub-dataset 1 used as test dataset; 0.00063 for sub-dataset 2; 0.00061 for sub-dataset 3; and 0.00037 for sub-dataset 4. The mean squared reconstruction error obtained from the trained PCA model on the

test dataset using the first three components was 0.00440 for sub-dataset 1 used as test dataset; 0.00082 for sub-dataset 2; 0.00099 for sub-dataset 3; and 0.00240 for sub-dataset 4.

The MSE results obtained from all the models on non-training data are shown in Table 6. Based on Table 6, it can be concluded that the MLP network provides accurate predictions for all the measurands and the differences in the MSE values between each sub-dataset are very small. Also, the residual values were calculated by the absolute difference between the measured deviations obtained from the comparator system and the predicted deviations. The residual values range from 0.4 μm to 6.3 μm for diameter, from 0.0 μm to 1.3 μm for true position, and from 0.0 μm to 2.7 μm for circularity. The average residual values for diameter, true position and circularity are 3.0 μm, 0.6 μm and 1.3 μm, respectively. Therefore, it can be concluded that feedforward neural networks perform well in predicting the end product quality deviations especially for true position and circularity.

The proposed method for inspection by exception is based on the following scheme to identify the parts that require inspection in order to assess whether or not they conform to design specifications. Given the predicted product health metric deviation vector $\hat{\boldsymbol{h}}_j$: the parts with $\hat{h}_{jp} < \underline{\tau}_j$, for $p = 1, 2, \ldots, \tilde{P}$, are conforming parts and do not require inspection because their predicted deviations are far from the specified tolerances; in order to account for the uncertainty associated with the model predictions, the parts with $\underline{\tau}_j \leq \hat{h}_{jp} \leq \overline{\tau}_j$ require inspection because their predicted deviations are close to the specified tolerances; and the parts with $\hat{h}_{jp} > \overline{\tau}_j$ are non-conforming parts with no requirements for inspection because their predicted deviations have greatly exceeded the tolerance specifications. Given the model predictions, $\hat{\boldsymbol{h}}_j$, and part tolerances (±0.0700 mm for diameter, 0.0100 mm for true position and 0.0500 mm for circularity), the non-conforming parts are the parts: 1, 10, and 23. The predicted diameter deviations for these three parts are: 0.0766 mm, 0.0713 mm, and 0.0794 mm, respectively. Accounting for ±6 μm, ±2 μm and ±3 μm uncertainty associated with the model predictions, for diameter deviation, true position, and circularity, respectively, and thus, given that $\underline{\tau}_j = 0.0640$ mm, where $j = 1$ for diameter deviation, $\underline{\tau}_j = 0.0080$ mm, where $j = 2$ for true position, and $\underline{\tau}_j = 0.0470$ mm, where $j = 3$ for circularity, and, $\overline{\tau}_j = 0.0760$ mm, where $j = 1$ for diameter deviation, $\overline{\tau}_j = 0.0120$ mm, where $j = 2$ for true position, and $\overline{\tau}_j = 0.0530$ mm, where $j = 3$ for circularity, only the part 10 requires inspection because its predicted diameter deviation is within $\underline{\tau}_1$ and $\overline{\tau}_1$. Given these bounds, the proposed method based on PCA-based neural networks can reduce inspection volume from sixteen parts to just one part.

For comparison, linear regression models were also developed. Figs. 9–11 show the normal probability plots of the residuals of the fitted linear regression models. The MSE results on non-training data are shown in Table 7. Figs. 12–14 are the bar graphs of residuals calculated by the absolute difference between the measured deviations obtained from the comparator measurement system and the predicted deviations obtained from the neural network and the linear model. Linear regression models are among the most fundamental and widely used tools for many modelling problems due to their simplicity, interpretability and performance in low-data regimes. However, they may be inadequate as models for nonlinear problems and, as a result, they may not fit the data as well as neural networks. Many advanced regression modelling methods including neural networks can be considered as extensions of linear regression modelling [34]. Compared to linear regression, the neural network model provides more accurate predictions. For the linear regression model, the residual values range from 0.2 μm to 12.9 μm for diameter, from 0.1 μm to 1.8 μm for true position, and from 0.2 μm to 5.2 μm for circularity. The average residual values for the regression model for diameter, true position and circularity are 5.0 μm, 0.8 μm, and 2.3 μm, respectively. The predicted diameter deviations from the regression model for parts 1, 10 and 23 are: 0.0642 mm, 0.0583 mm, and 0.0857 mm, respectively, and thus, the linear regression model fails to achieve inspection by exception. Tables 8 and 9 show the predicted deviations from the neural network model and linear regression model, respectively.

Table 6
4-fold cross-validation for neural network model.

Sub-dataset	Diameter deviation MSE (mm^2)	True position MSE (mm^2)	Circularity MSE (mm^2)
1	1.21×10^{-5}	5.48×10^{-7}	1.16×10^{-6}
2	1.25×10^{-5}	7.80×10^{-7}	4.54×10^{-6}
3	2.26×10^{-5}	3.99×10^{-7}	3.24×10^{-6}
4	5.93×10^{-6}	3.55×10^{-7}	5.10×10^{-7}

Table 7
4-fold cross-validation for linear regression model.

Sub-dataset	Diameter deviation MSE (mm^2)	True position MSE (mm^2)	Circularity MSE (mm^2)
1	2.68×10^{-5}	4.90×10^{-7}	8.92×10^{-6}
2	1.18×10^{-5}	1.22×10^{-6}	1.10×10^{-5}
3	4.87×10^{-5}	1.00×10^{-6}	5.39×10^{-6}
4	5.58×10^{-5}	7.76×10^{-7}	5.47×10^{-6}

Table 8
Predicted deviations from neural network model.

Parts	Diameter deviation (mm)	True position (mm)	Circularity(mm)
2	0.0533	0.0043	0.0364
23	0.0794	0.0056	0.0400
22	0.0523	0.0038	0.0355
21	0.0543	0.0017	0.0378
12	0.0517	0.0029	0.0382
13	0.0570	0.0054	0.0383
24	0.0494	0.0029	0.0383
7	0.0499	0.0053	0.0378
19	0.0575	0.0061	0.0382
6	0.0541	0.0021	0.0367
9	0.0554	0.0035	0.0372
10	0.0713	0.0057	0.0447
1	0.0766	0.0073	0.0416
16	0.0457	0.0058	0.0370
15	0.0545	0.0036	0.0368
18	0.0537	0.0037	0.0357

Table 9
Predicted deviations from linear regression model.

Parts	Diameter deviation (mm)	True position (mm)	Circularity (mm)
2	0.0571	0.0038	0.0378
23	0.0857	0.0064	0.0462
22	0.0580	0.0052	0.0390
21	0.0547	0.0026	0.0380
12	0.0553	0.0031	0.0360
13	0.0605	0.0057	0.0392
24	0.0540	0.0031	0.0358
7	0.0593	0.0056	0.0388
19	0.0573	0.0055	0.0383
6	0.0577	0.0031	0.0378
9	0.0539	0.0033	0.0373
10	0.0583	0.0043	0.0381
1	0.0642	0.0067	0.0385
16	0.0545	0.0049	0.0380
15	0.0518	0.0032	0.0367
18	0.0566	0.0036	0.0372

6.2.4. Conformance and non-conformance probabilities

The probability that a random variable H is no greater than a specified value T is given by the integral of the PDF between $-\infty$

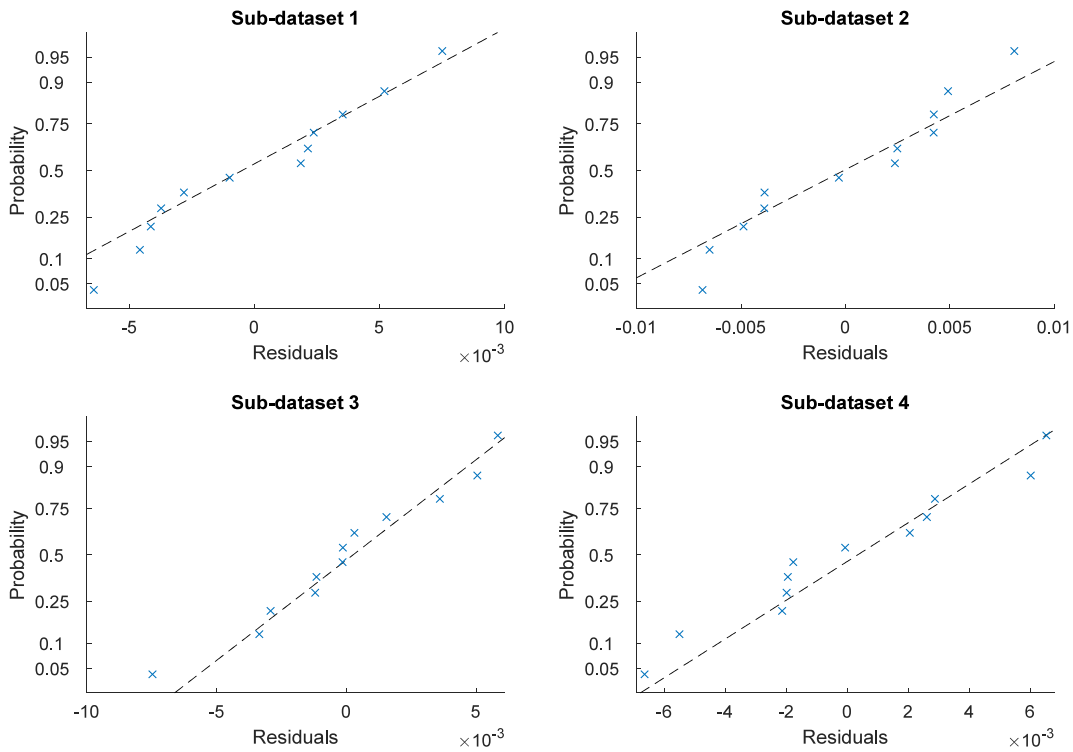

Fig. 9. Normal probability plots of residuals for diameter deviation.

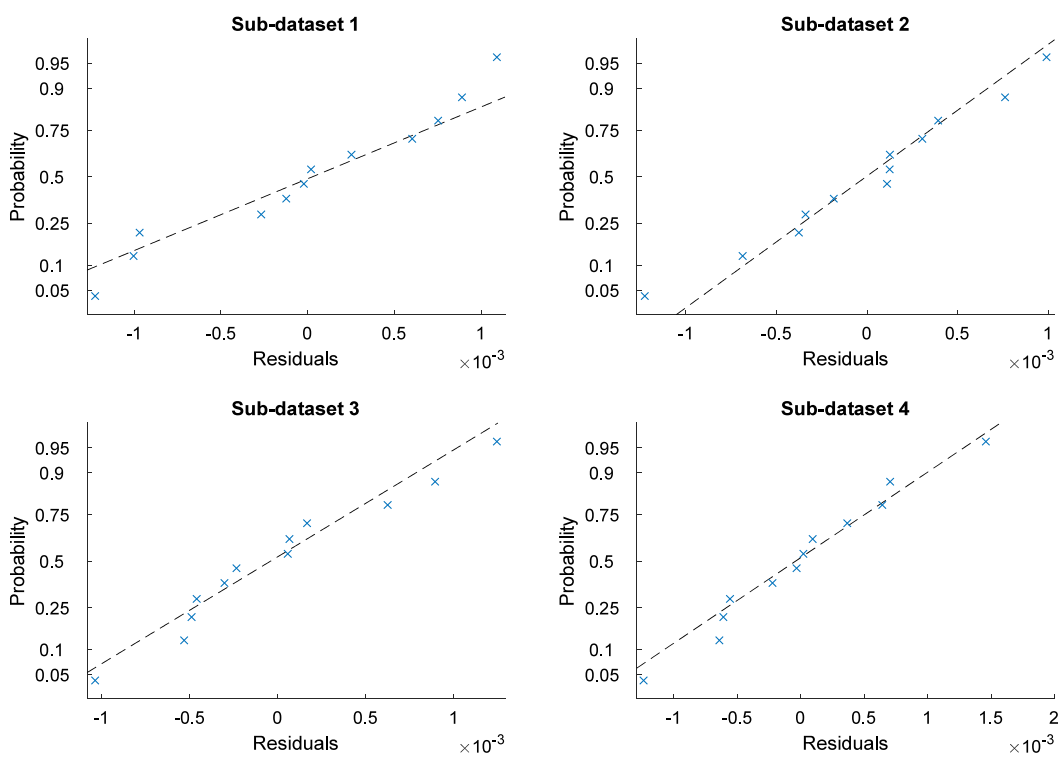

Fig. 10. Normal probability plots of residuals for true position.

and T and represented by the Cumulative Distribution Function (CDF):

$$G(T) = P(H \leq T) = \int_{-\infty}^{T} g(h)\, dh, \quad T \in \Re, \tag{16}$$

with $\lim_{T\to-\infty} G(T) = 0$ and $\lim_{T\to\infty} G(T) = 1$. Therefore, the CDF can be used to compute the conformance probability:

$$p_c = P(T_L \leq H \leq T_U) = G(T_U) - G(T_L) = \int_{T_L}^{T_U} g(h)\, dh. \tag{17}$$

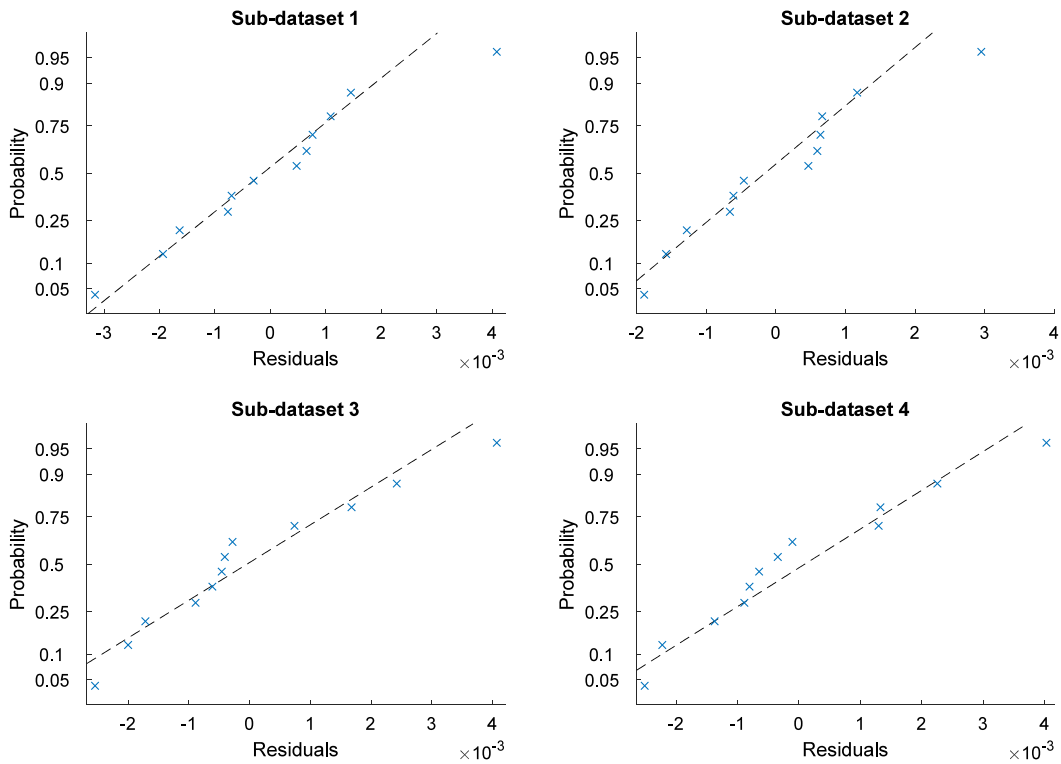

Fig. 11. Normal probability plots of residuals for circularity.

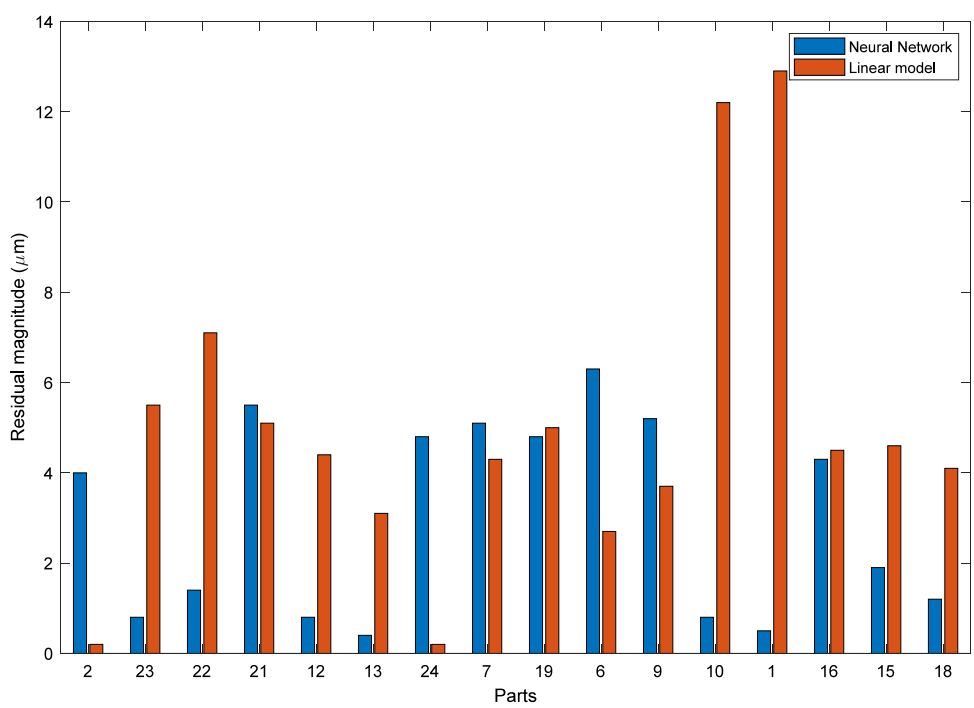

Fig. 12. Residual magnitude for diameter deviation for all parts evaluated by cross validation.

Given the tolerance limits T_L and T_U for diameter, the best estimate h_{1p}, and associated standard uncertainty $u(h_{1p})$, the conformance probability is $p_c(h_{1p}) = \Phi\left(\frac{T_U - h_{1p}}{u(h_{1p})}\right) - \Phi\left(\frac{T_L - h_{1p}}{u(h_{1p})}\right)$, where Φ is the CDF of the standard normal random variable with zero mean and unit variance. For example, $p_c(h_{1p}) = 0.3$ for $p = 10$ (part 10) given that the tolerance limits are -0.07 mm and 0.07 mm, the mean is 0.0705 mm, and the standard uncertainty is 1 μm. The non-conformance probability is $\overline{p}_c = 1 - p_c$. Similarly, the conformance and non-conformance probabilities can be obtained from the prediction results $\hat{h}_{jp}$ by calculating an estimate of the variance from the residuals. However, for true position and circularity, one-sided tolerance intervals with a single upper tolerance limit are used and thus, $p_c(h_{jp}) = \Phi\left(\frac{T_U - h_{jp}}{u(h_{jp})}\right)$ for $j = 2$ and $j = 3$. Note that, although the (diameter) deviations

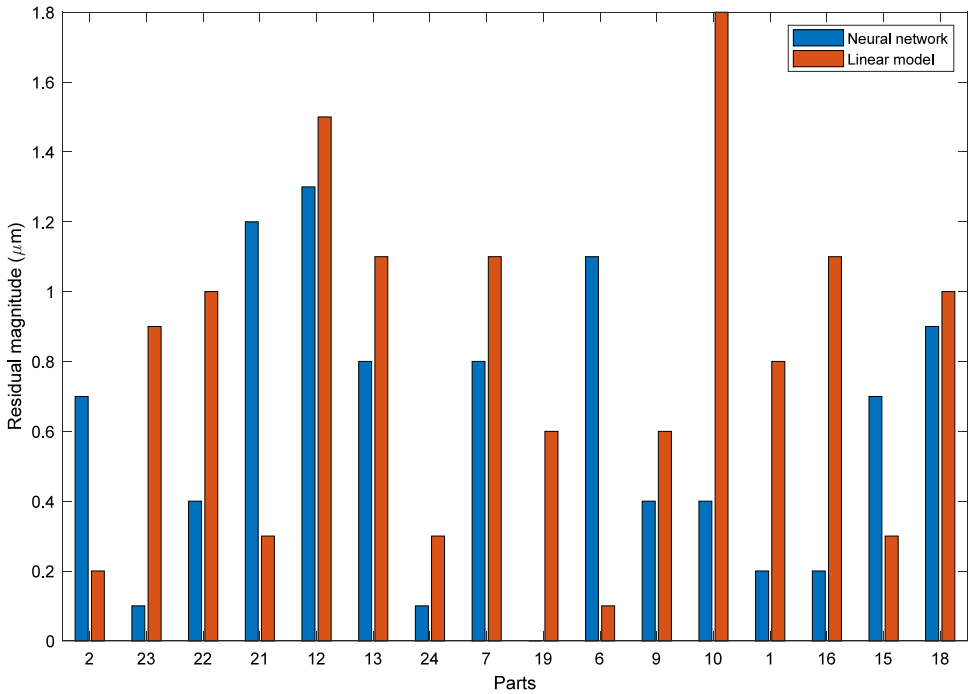

Fig. 13. Residual magnitude for true position for all parts evaluated by cross validation.

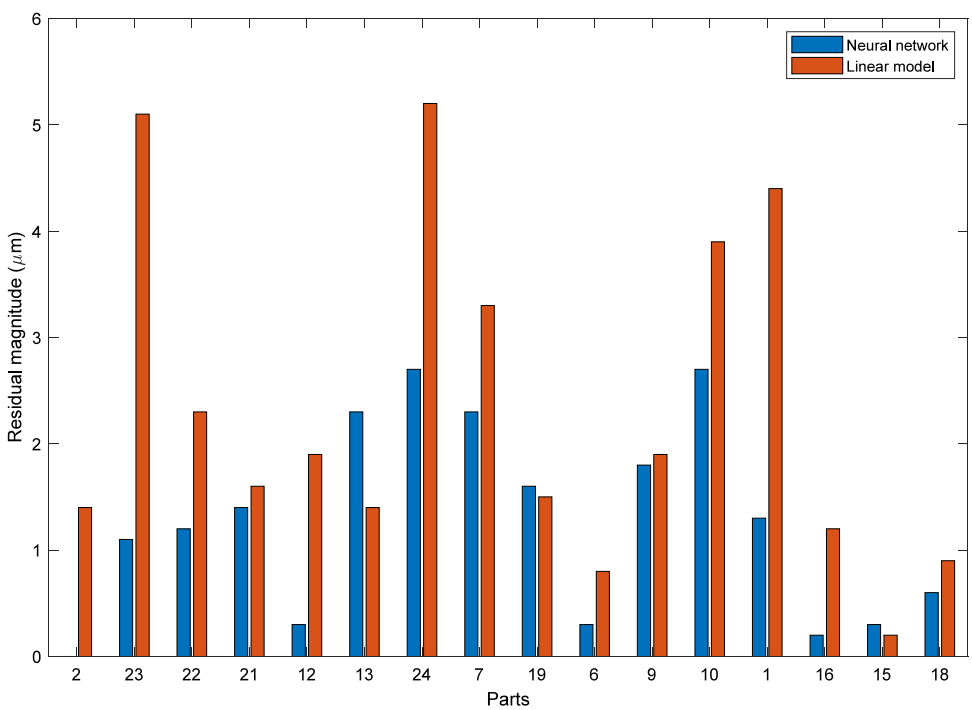

Fig. 14. Residual magnitude for circularity for all parts evaluated by cross validation.

have been obtained as absolute values, two-sided tolerance intervals with lower and upper tolerance limits are still required to obtain the conformance and non-conformance probabilities for diameter deviations. Finally, signed deviations must be used for non-symmetrical two-sided tolerance intervals.

7. Summary and concluding remarks

With the proliferation of in-process metrology data over the years, there has been an increase of interest in machine learning techniques to develop more efficient monitoring and control strategies for intelligent/smart manufacturing. The choice of data analytics and machine learning algorithms used to discover automatically useful patterns and trends in collected data or perform accurate predictions depends on many factors such as the cost and practicality in collecting the class labels. This paper presented a new method for what the authors have called 'inspection by exception'. The proposed method is based on multistage manufacturing data and machine learning techniques, including both unsupervised and supervised learning algorithms. The input data samples included material conditions, tempering temperature,

and vibration and force signals of the metal cutting process. PCA was used for dimensionality reduction. An unsupervised learning approach was developed to partition the manufactured products into two groups using the first two principal components of the normalized dataset and the FCM clustering algorithm. The normalized Euclidean distances between the principal components and cluster centres were then computed to identify the conforming parts, the non-conforming parts, and the parts that require inspection in order to determine whether or not they conform to specifications. The unsupervised learning approach reduced significantly the requirement for inspection, and, in particular, based on the 2-fold cross validation results, it was demonstrated that the volume of inspections were reduced from seventeen parts to only three parts.

A supervised learning approach based on neural networks was also developed to provide deviation vectors for product health using the first three principal components of the normalized dataset. The output data used for training the supervised model were obtained from an automated comparator system. An experimental methodology to evaluate the uncertainty associated with comparator measurement was presented to provide confidence in the measurement results and enable the calculation of conformance and non-conformance probabilities. The predicted product health metric deviation vectors compared well with the measured ones especially for true position and circularity. Regression models were also developed and compared with the neural network models. The results showed that the neural network models outperform the regression models. Based on the 4-fold cross validation results from the supervised learning approach, the volume of inspections were reduced from sixteen parts to just one part. The proposed method for inspection by exception accounts for model uncertainties and therefore, it can comply with applications that suffer from tight part tolerances. Determining the parts that require inspection for conformance assessment in an unsupervised manner has the advantage of lower cost since class labels are not required for training the machine learning algorithm. However, obtaining the health condition of the parts require supervised models.

CRediT authorship contribution statement

Moschos Papananias: Implemented the proposed approach and prepared the manuscript. **Thomas E. McLeay:** Conceived the proposed approach and reviewed the manuscript. **Olusayo Obajemu:** Reviewed the manuscript. **Mahdi Mahfouf:** Reviewed the manuscript. **Visakan Kadirkamanathan:** Supervised the implementation and reviewed the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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] M.P. Groover, Fundamentals of Modern Manufacturing: Materials Processes, and Systems, John Wiley & Sons, 2007.
[2] J. Shi, Stream of Variation Modeling and Analysis for Multistage Manufacturing Processes, CRC press, 2006.
[3] R.J. Hocken, P.H. Pereira, Coordinate Measuring Machines and Systems, CRC Press, 2016.
[4] P. Pereira, R. Hocken, Characterization and compensation of dynamic errors of a scanning coordinate measuring machine, Precis. Eng. 31 (1) (2007) 22–32.
[5] J.V. Abellan-Nebot, F.R. Subirón, A review of machining monitoring systems based on artificial intelligence process models, Int. J. Adv. Manuf. Technol. 47 (1–4) (2010) 237–257.
[6] T.E. McLeay, Unsupervised Monitoring of Machining Processes, University of Sheffield, 2016.
[7] D. Ceglarek, et al., Time-based competition in multistage manufacturing: Stream-of-variation analysis (SOVA) methodology, Int. J. Flexible Manuf. Syst. 16 (1) (2004) 11–44.
[8] F. Yang, S. Jin, Z. Li, A comprehensive study of linear variation propagation modeling methods for multistage machining processes, Int. J. Adv. Manuf. Technol. 90 (5–8) (2017) 2139–2151.
[9] J. Shi, in: J. Baillieul, T. Samad (Eds.), Stream of Variations Analysis, in Encyclopedia of Systems and Control, Springer London, London, 2014, pp. 1–6.
[10] J.-P. Loose, S. Zhou, D. Ceglarek, Kinematic analysis of dimensional variation propagation for multistage machining processes with general fixture layouts, IEEE Trans. Autom. Sci. Eng. 4 (2) (2007) 141–152.
[11] A. Bazdar, R.B. Kazemzadeh, S.T.A. Niaki, Fault diagnosis within multistage machining processes using linear discriminant analysis: a case study in automotive industry, Qual. Technol. Quant. Manage. 14 (2) (2017) 129–141.
[12] S. Du, et al., Three-dimensional variation propagation modeling for multistage turning process of rotary workpieces, Comput. Ind. Eng. 82 (2015) 41–53.
[13] K. Wang, et al., State space modelling of variation propagation in multistage machining processes for variable stiffness structure workpieces, Int. J. Prod. Res. (2020) 1–20.
[14] G. Wang, et al., A generative neural network model for the quality prediction of work in progress products, Appl. Soft Comput. 85 (2019) 105683.
[15] L. Li, K. Ota, M. Dong, Deep learning for smart industry: Efficient manufacture inspection system with fog computing, IEEE Trans. Ind. Inf. 14 (10) (2018) 4665–4673.
[16] M. Papananias, et al., A Bayesian framework to estimate part quality and associated uncertainties in multistage manufacturing, Comput. Ind. 105 (2019) 35–47.
[17] D.R. Salgado, et al., In-process surface roughness prediction system using cutting vibrations in turning, Int. J. Adv. Manuf. Technol. 43 (1–2) (2009) 40–51.
[18] T. Özel, Y. Karpat, Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks, Int. J. Mach. Tools Manuf. 45 (4–5) (2005) 467–479.
[19] P.B. Huang, An intelligent neural-fuzzy model for an in-process surface roughness monitoring system in end milling operations, J. Intell. Manuf. 27 (3) (2016) 689–700.
[20] P. Kovac, et al., Application of fuzzy logic and regression analysis for modeling surface roughness in face milliing, J. Intell. Manuf. 24 (4) (2013) 755–762.
[21] C. Han, M. Luo, D. Zhang, Optimization of varying-parameter drilling for multi-hole parts using metaheuristic algorithm coupled with self-adaptive penalty method, Appl. Soft Comput. (2020) 106489.
[22] I. Jolliffe, Principal component analysis, in: International Encyclopedia of Statistical Science, Springer, 2011, pp. 1094–1096.
[23] R.B. Bapat, Linear Algebra and Linear Models, Springer Science & Business Media, 2012.
[24] J.C. Bezdek, R. Ehrlich, W. Full, FCM: The fuzzy c-means clustering algorithm, Comput. Geosci. 10 (2–3) (1984) 191–203.
[25] S.S. Haykin, Neural Networks and Learning Machines, Vol. 3, Pearson Upper Saddle River, NJ, USA, 2009.
[26] M. Papananias, et al., Uncertainty evaluation associated with versatile automated gauging influenced by process variations through design of experiments approach, Precis. Eng. 49 (2017) 440–455.
[27] M. Papananias, et al., Evaluation of automated flexible gauge performance using experimental designs, in: Laser Metrology and Machine Performance XII, LAMDAMAP, euspen, Wotton-under-Edge, UK, 2017.
[28] M. Papananias, et al., An intelligent metrology informatics system based on neural networks for multistage manufacturing processes, Procedia CIRP 82 (2019) 444–449.
[29] A. Forbes, A. Mengot, K. Jonas, Uncertainty associated with coordinate measurement in comparator mode, in: Laser Metrology and Machine Performance XI, LAMDAMAP, euspen, Huddersfield, UK, 2015.
[30] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of measurement data — Guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology, JCGM 100: 2008, GUM 1995 with minor corrections.
[31] I. Lira, Evaluating the Measurement Uncertainty: Fundamentals and Practical Guidance, IOP, 2002.

[32] BS EN ISO 15530-3: 2011: Geometrical product specifications (GPS). Coordinate measuring machines (CMM). Technique for determining the uncertainty of measurement. Use of calibrated workpieces or measurement standards. British Standards Institute.

[33] M. Riedmiller, H. Braun, A direct adaptive method for faster backpropagation learning: The RPROP algorithm, IEEE International Conference on Neural Networks, San Francisco, CA, USA, 1993.

[34] S. Weisberg, Applied Linear Regression, Vol. 528, John Wiley & Sons, 2005.