



This is a repository copy of *A Bayesian information fusion approach for end product quality estimation using machine learning and on-machine probing*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/186737/>

Version: Accepted Version

Article:

Papananias, M., McLeay, T.E., Mahfouf, M. orcid.org/0000-0002-7349-5396 et al. (1 more author) (2022) A Bayesian information fusion approach for end product quality estimation using machine learning and on-machine probing. *Journal of Manufacturing Processes*, 76. pp. 475-485. ISSN 1526-6125

<https://doi.org/10.1016/j.jmapro.2022.01.020>

© 2022 Published by Elsevier Ltd on behalf of The Society of Manufacturing Engineers. This is an author produced version of a paper subsequently published in *Journal of Manufacturing Processes*. Uploaded in accordance with the publisher's self-archiving policy. Article available under the terms of the CC-BY-NC-ND licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. 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.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

A Bayesian information fusion approach for end product quality estimation using machine learning and on-machine probing

Moschos Papananias^{a*}, Thomas E McLeay^b, Mahdi Mahfouf^a, Visakan Kadirkamanathan^a

^aDepartment of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

^bSandvik Coromant, Mossvågen 10, Sandviken 811 34, Sweden

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

Abstract

There is an increasing demand for manufacturing processes to improve product quality and production rates while keeping the costs to a minimum. The quality of the products is influenced by several sources of errors introduced during the series of manufacturing operations. These errors accumulate over these multiple stages of manufacturing. Therefore, monitoring systems for product health utilising data and information from different sources and manufacturing stages is a key factor to meet these growing demands. This paper is concerned with the process of combining new measurement data or information with machine learning-based prediction information obtained as each product goes through a series of manufacturing steps to update the conditional probability distribution of the end product quality during manufacturing. A Bayesian approach is adopted in obtaining an updated posterior distribution of the end product quality given new information from subsequent measurements, and, in particular, On-Machine Probing (OMP). Following the steps of heat treatment, machining, and OMP, the posterior distribution of the previous step can be considered as the new prior distribution to obtain an updated posterior distribution of the product condition as new metrological information becomes available. It is demonstrated that the resulting posterior estimates can lead to more efficient product condition monitoring in multistage manufacturing.

Keywords: Bayesian Inference; Machine Learning; Information Fusion; Multistage Manufacturing Process (MMP); On-Machine Probing (OMP); Uncertainty of Measurement

1. Introduction

Manufacturing is concerned with transforming starting raw materials into finished parts or products designed usually with exceptionally tight tolerances. One of the most important manufacturing methods is the Computer Numerical Control (CNC) machining in which unwanted material is removed from a

workpiece in the form of small chips by means of a rotary cutting tool that moves along certain multiple axes as indicated by a customised computer program [1], [2]. In machining, part accuracy is affected by many sources of errors of varying magnitude, such as geometric and kinematic errors, thermal errors, cutting force-induced errors, fixturing errors, and tool wear [3], [4]. In addition, a workpiece to be machined may have already gone through other processing stages, such as forming and heat treatment. Therefore, in Multistage Manufacturing Processes (MMPs), product quality variations are contributed by the errors generated at the current manufacturing stage, as well as the accumulated errors transmitted from preceding stages [5]–[7].

The manufacturing industry is currently undergoing a significant transformation towards the concept of smart manufacturing or Industry 4.0 concerned with a new generation of manufacturing processes characterized with autonomy and intelligence based on Cyber-Physical Systems (CPSs) [8], [9]. The factory of the future will operate with manufacturing equipment and systems capable of being self-optimized and communicating with each other for making optimal decisions for example, in the event of producing out-of-specifications products. Manufacturing processes are increasingly equipped with various sensors and data acquisition devices to gather data as the product is manufactured for process and product health monitoring and control [10]. Continuous condition monitoring and control of the process and product being manufactured make production more flexible with greater manufacturing efficiency and productivity. During recent years, a significant interest has been devoted to multisensor data fusion in dimensional metrology [11] for combining data from multiple sources, such as dimensional inspection data from a Coordinate Measuring Machine (CMM) and a structured line scanner [12]. Multisensor data fusion is a multidisciplinary field of interest that aims to overcome the limitations of individual data acquisition devices and reduce the uncertainty of information estimates. It is worth distinguishing between data fusion and information fusion though these terms are sometimes used interchangeably. The former refers to the combination of data obtained directly from multiple sensors while the latter refers to the combination of already processed data and information from sensors, technical reports, models, etc. [13].

Over the years, several methods have been proposed to develop monitoring systems based on machine learning process models and in-process metrology data for observing product health, especially surface quality characteristics such as surface roughness [14]. However, published work on intelligent condition monitoring systems for dimensional metrology characteristics is limited [15]. Also, most predictive monitoring methods are based on mapping single stage manufacturing data to product quality characteristics, though manufacturing processes typically involve multiple production stages, and do not take into account further metrological information obtained at a later manufacturing stage. Because of the complexity and diversity of manufacturing processes, developing a reliable and robust condition monitoring system that has minimal implementation and maintenance costs requires efficient and flexible modelling techniques. The use of Bayesian methods has greatly increased during recent years in many applications including manufacturing process monitoring and control [16], [17]. Bayesian methods make use of probability distributions to quantify uncertainty in statistical inferences and can be seen as being equivalent to combining information from multiple sources [18]–[20]. In a Bayesian formulation, probability distributions are updated as more information becomes available. This paves the road to

monitor the health of the process and product using statistical and machine learning techniques, as it goes through multiple manufacturing steps, and update our predictions about the product quality in the light of subsequent measurements through Bayes' theorem. This increases the reliability of metrological information estimates through the different stages. For example, in situ dimensional inspection of finished or semi-finished parts on the machine with a Touch-Trigger Probe (TTP) enables the evaluation of dimensional metrological information about the product in a single setup and can be considered as an additional source of information for Bayesian data or information fusion. On-Machine Probing (OMP) using TTPs can reduce unnecessary downtime, re-work, scrap, and post-process inspection [21]. However, the use of a CNC machine tool as a Coordinate Measuring System (CMS) also has some drawbacks. In particular, OMP is fraught with the same error sources as CMMs, and, in addition, it cannot detect machine tool error-induced deviations. CMM measurement results are subject to many multivariate influence factors, such as geometric and kinematic errors, probing system errors, and environmental effects [22], [23]. Therefore, the evaluation of the uncertainty associated with CMM measurement and OMP is not straightforward. In this work, a CMS based on parallel kinematic configuration and operating in comparator mode (Equator gauging system) was used for post-process inspection to obtain the dimensional metrology characteristics of machined parts. Operating a CMS in comparator mode has the advantage of obtaining measurements that are devoid of constant systematic effects associated with the measurement system [22], [24].

The growing demand for improved product quality and production rates with reduced inspections have contributed significantly to the development of advanced process monitoring and control systems. However, the performance of machining processes, such as turning, drilling, milling, and grinding, depends on several parameters, including machine configuration, machining parameters and tool path trajectory, cutting tool type and wear, workpiece material and fixturing, process dynamics, etc. Therefore, the systems require training with a particular manufacturing method or fault type to provide an accurate prediction and then updating this prediction, given that new information is obtained. The research of this article applies a Bayesian fusion method to provide an improved product health parameter estimate using probabilistic machine learning and OMP. Unlike other studies in this domain, additional on-machine **inspection** data are utilized, so that the system can update its predictions as the product is manufactured when new metrology data become available. In the case presented, metrology data are obtained from a MMP consisting mainly of heat treatment, machining, and dimensional inspections.

The purpose of this paper is to present a novel approach to update the conditional probability distribution of the end product quality, by using Bayesian information fusion of machine learning based estimation and subsequent measurements, such as OMP. The advantages of the proposed approach to product health monitoring for MMPs naturally arise from the sequential nature of Bayes' theorem for updating posterior distributions. The proposed method is validated using experimental data from a multistage manufacturing case study (see Figure 1). Section 2 presents a detailed review of literature relating to dimensional metrology and process monitoring and control methods. Section 3 describes the product health metric deviation matrix and the probabilistic model. Section 4 validates the model using data from an experimental case study concerned with the manufacture of steel bearing housing parts [25]–[27].

Section 5 presents the proposed Bayesian information fusion approach and the results obtained with this method. Concluding remarks are given in Section 6.

2. Related literature

Dimensional metrology is concerned with the measurement of geometric features of a manufactured part in order to determine whether or not the part conforms to its geometric tolerance specifications (form, orientation, profile, runout, size, and location). Dimensional measurements can be obtained by a variety of methods including both manual inspection methods, such as hard gauging and Articulated Arm Coordinate Measuring Machine (AACMM) measurement, and automated inspection methods, such as OMP, CMM measurement, and flexible gauging [28]. Manual inspection methods are prone to a non-predictable error source, the operator, affecting the repeatability, reproducibility, and part throughput, and usually lead to high measurement uncertainties. Coordinate metrology, particularly the use of CMMs has become vital for industrial dimensional metrology because of their efficiency, flexibility and accuracy. However, providing valid uncertainty statements associated with a particular CMM measurement task requires significant efforts as CMM measurement results are subject to a large range of influence factors including both random and systematic effects [22]. CMSs can also be used to perform coordinate measurements in a gauging/comparator mode in which dimensional measurements of a workpiece are compared with those of a master artefact nominally of the same geometry. Coordinate measurements made in a comparator mode benefit from the fact that many of the systematic effects associated with the CMS cancel out and thus need not be modelled when evaluating the uncertainties. However, the traceability path associated with comparative coordinate measurement is not as well defined because such measurement results originate from indirect/relative measurement. In addition, the measurement uncertainty for a given workpiece measured using a comparator measurement system will always inherit uncertainty from the calibration of the master artefact, but this uncertainty contributor is usually not difficult to quantify. Process variations, such as part misalignments from rotation between master and measure coordinate reference frames, are also possible uncertainty contributors for comparative coordinate measurement, particularly when using a non-repeatable fixturing setup [28]. Furthermore, establishing a master artefact for comparator measurement may be not straightforward and usually requires calibrating a manufactured part produced close to drawing nominals on an accurate CMM.

Traditional part quality assessment techniques are usually based on manually-operated measurement instruments and CMMs that can potentially create production bottlenecks limiting production rate. Therefore, in recent years, there has been a drive towards process and product health monitoring strategies based on machine learning and live captured sensor signals in order to make timely decisions while minimising the volume of non-value adding processes, such as dimensional inspection. In particular, Industry 4.0 has become the new trend in the manufacturing industry and much research has focused on mapping process parameters and monitoring data to product quality characteristics, such as surface roughness, using machine learning process models. Özel and Karpaz [29] used Artificial Neural Networks (ANNs) to predict both surface roughness and tool flank wear in finish dry hard turning using material hardness, cutting speed, feed rate, axial cutting length, and force data. Plaza et al. [30] analyzed different feature extraction methods to optimize surface finish monitoring in CNC turning using vibration signals

obtained by a single low-cost accelerometer sensor. Salgado et al. [31] developed Least Squares Support Vector Machines (LS-SVMs) to predict surface roughness in turning using cutting parameters (feed rate, cutting speed, and depth of cut), tool geometry parameters (nose radius and nose angle), and vibration data. Huang [32] developed a neural-fuzzy monitoring system for end-milling operations to predict surface roughness using spindle speed, feed rate, depth of cut, the average resultant peak force, and the absolute average force. Kovac et al. [33] applied fuzzy logic and regression analysis for modelling surface roughness in dry face milling using machining conditions including cutting speed, feed rate, depth of cut, and width of flank wear land. Bolar et al. [34] performed a full factorial design to investigate the influence of feed per tooth, tool diameter, and axial and radial depths of cut on cutting forces and surface roughness during machining of thin-wall parts and developed second order regression models for the prediction of both measurands given the studied process parameters. Han et al. [35] proposed a varying-parameter drilling method to improve manufacturing efficiency in successive drilling operations and ultimately increase hole surface roughness quality for multi-hole parts made of difficult-to-cut materials. The hole surface roughness was predicted by Radial Basis Function (RBF) networks using spindle speed, feed rate, crater wear, flank wear, outer corner wear, thrust force and torque. Moore et al. [36] proposed a machine learning-based machine and process monitoring system for milling using vibration and power signals. Correa et al. [37] applied Bayesian networks and ANNs to predict surface roughness in high-speed milling using various features, including workpiece geometry, material hardness, machining parameters, and cutting forces. They showed that Bayesian networks are easier to interpret than ANNs and performed better in this classification problem.

Beyca et al. [38] developed a Bayesian Dirichlet Process (DP) multisensor fusion decision theoretic approach to detect abnormal process drifts in ultraprecision machining by integrating multiple in situ sensor signals, such as force, vibration, and Acoustic Emission (AE). The results showed that their approach can classify ultraprecision machining process drifts much more accurately than conventional classification methods, such as ANNs and SVMs. Karandikar et al. [39] implemented a Bayesian learning method for stability lobe identification using milling test results and presented an adaptive experimental strategy to identify the optimal combination of parameters, which maximise material removal rate. Wang et al. [40] proposed a multisensor fusion method of vision and sound to monitor in-process grinding material removal rate. They conducted belt grinding experiments using different parameters and derived a predictive model for material removal rate monitoring based on the optimal feature subsets and an improved light gradient boosting machine algorithm. Nazir and Shao [41] proposed an online tool condition monitoring system for ultrasonic metal welding using sensor fusion and machine learning techniques. They tested a variety of classification models using experimental data and concluded that displacement and AE sensor signals are more useful in predicting tool conditions than power and sound signals. Atoui et al. [42] presented a probabilistic framework for system monitoring based on Bayesian networks, but their approach was tested with a simulation of a water heater process. Zhao et al. [43] derived an algorithm based on a linear state-space model for sensor monitoring, which estimates the probability distributions of measurement noise covariance and state variable simultaneously, assuming that all the sensors are uncorrelated with each other. The joint posterior distribution was approximated by two independent proposal distributions under the variational Bayesian inference framework, but the algorithm was tested with a numerical simulation and a quadruple water tank experiment. Du et al. [44] developed a Bayesian monitoring method based on a linear state-space model to estimate the process control parameters and establish the control limits of the cause-selecting chart in the ramp-up phase of a

MMP. Tran et al. proposed two one-sided Shewhart-type charts to monitor the ratio of two normal random variables for a finite horizon production as there are numerous situations where the production run is finite, but they considered simulation data from the food industry for the quality control problem. Riaz et al. [45] presented Bayesian posterior predictive exponentially weighted moving average control charts under different loss functions for small to moderate process mean shift detection, but the applications considered for validation were not in the manufacturing sector. For a lab-scale distillation column and the Tennessee Eastman (ET) industrial challenge problem, Ghosh et al. [46] used multiple heterogeneous fault detection and identification methods and fused their results to overcome the limitations of each method when used separately. Zhang and Ge [47] designed a fusion system by combining results of various methods for fault detection and identification in industrial processes. They used the Dempster-Shafer evidence theory to combine decisions generated from different models and a resampling strategy as a data pre-processing step to enhance the diversity performance of the fusion system. The case study included the ET process. In summary, Bayesian and machine learning methods via sensor and decision fusion techniques had been applied to specific manufacturing stages and to other industrial processes involving a variety of formulations ranging from classification to dynamic fault estimation, but homogeneously within the methods.

To the best of our knowledge, no research has been done regarding the fusion of in-process inspection data and in-process monitoring data for dimensional product health monitoring. The overwhelming majority of published research works have proposed machine learning-based methods to identify the finish-machined part condition, particularly surface metrology characteristics, by monitoring only the machining process, and do not enhance their predictions when new information becomes available. The aim of this article is to fill this gap by introducing a multisensor fusion method to predict an improved product health parameter estimate. First, this research work develops an intelligent, dimensional product health monitoring system, which learns from in-process metrology data obtained from multiple different manufacturing stages and sources, to predict the end product quality. Second, it develops a Bayesian information fusion algorithm to update this prediction as the product is manufactured, given new information from subsequent measurements, such as OMP. A Bayesian updating procedure is adopted in combining the information obtained from machine learning with the new information obtained from OMP. It is tested on a real industrial case study used to manufacture bearing housing parts made by EN24T steel. The manufacturing process involves heat treatment, grinding, hardness testing, machining, in-process inspection, and post-process inspection.

3. Product health metric deviation matrix and the probabilistic model

The concepts of basic probability and random variables that are used in this paper are first given. Let Y be a continuous random variable with Probability Density Function (PDF) $g(y)$. This function must: i) be nonnegative, i.e., $g(y) \geq 0$ for all y , and ii) have unit area, i.e., $\int_{-\infty}^{\infty} g(y)dy = P(-\infty < Y < \infty) = 1$. These two criteria must be strictly satisfied in order for a function to qualify as a PDF. The probability that the random variable Y is less than or equal to a value y is given by the Cumulative Distribution Function (CDF). The CDF $G(y)$ for the random variable Y is defined by $G(y) = P(Y \leq y)$, $-\infty < y < \infty$, and satisfies the marginal conditions $\lim_{y \rightarrow -\infty} G(y) = 0$ and $\lim_{y \rightarrow \infty} G(y) = 1$. The expectation $E(Y)$ of the

random variable Y with PDF $g(\boldsymbol{y})$ is $E(Y) = \mu_Y = \int_{-\infty}^{\infty} \boldsymbol{y}g(\boldsymbol{y})d\boldsymbol{y}$, the variance is $V(Y) = \sigma_Y^2 = E \left[(Y - E(Y))^2 \right] = \int_{-\infty}^{\infty} (\boldsymbol{y} - \mu_Y)^2 g(\boldsymbol{y})d\boldsymbol{y}$, and the standard deviation σ_Y is the positive square root of the variance, thus $\sigma_Y = \sqrt{V(Y)}$ [48].

In this paper, the product health information is represented through a metric deviation matrix. Suppose independent observations $\boldsymbol{y} = (\boldsymbol{y}_1, \dots, \boldsymbol{y}_m)^T$ of the random variable Y are available from a CMS, such as a CMM or a comparator measurement system. More generally, the measurement \boldsymbol{y}_i , for $i = 1, \dots, m$, is regarded as an observation of the random variable Y_i that has a PDF identical to that of Y . Consider now Y_{ij} independent random variables, for $i = 1, \dots, m$ and $j = 1, \dots, n$, and let \boldsymbol{H} be the corresponding product health metric deviation matrix:

$$\boldsymbol{H} = (h_{ij}) = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{m1} & h_{m2} & \cdots & h_{mn} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{y}}_1 - \boldsymbol{y}_{11} & \tilde{\boldsymbol{y}}_2 - \boldsymbol{y}_{12} & \cdots & \tilde{\boldsymbol{y}}_n - \boldsymbol{y}_{1n} \\ \tilde{\boldsymbol{y}}_1 - \boldsymbol{y}_{21} & \tilde{\boldsymbol{y}}_2 - \boldsymbol{y}_{22} & \cdots & \tilde{\boldsymbol{y}}_n - \boldsymbol{y}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\boldsymbol{y}}_1 - \boldsymbol{y}_{m1} & \tilde{\boldsymbol{y}}_2 - \boldsymbol{y}_{m2} & \cdots & \tilde{\boldsymbol{y}}_n - \boldsymbol{y}_{mn} \end{bmatrix} \in \mathcal{R}^{m \times n}, \quad (1)$$

where $h_{ij} = \tilde{\boldsymbol{y}}_j - \boldsymbol{y}_{ij}$ denotes the element located in the i th row and the j th column of the product health metric deviation matrix \boldsymbol{H} , $\tilde{\boldsymbol{y}}_j$ is the Computer-Aided Design (CAD) model based nominal value of the random variable or measurand Y_j , and \boldsymbol{y}_{ij} is the i th observation of the measurand Y_j . Given a tolerance specification T_j and a measured product health metric deviation h_{ij} , the product can be either accepted or rejected for the j th quality characteristic or, in some cases, it may be re-worked or re-measured with the same or more accurate measuring system.

The focus of this paper is to reduce inspections by firstly, creating a predictive model of the post-process inspection results from the in-process monitoring data. For each measurand j , consider a model of the form:

$$\boldsymbol{h} = \boldsymbol{X}\boldsymbol{a} + \boldsymbol{\epsilon}, \quad \boldsymbol{h} \in \mathcal{H}, \quad \mathcal{H} \sim N_p(\boldsymbol{X}\boldsymbol{a}, \sigma_{\boldsymbol{\epsilon}}^2 \boldsymbol{I}_p), \quad (2)$$

where $\boldsymbol{h} = (\boldsymbol{h}_1, \dots, \boldsymbol{h}_p)^T$ is the response variable obtained from a CMS operating in comparator mode with $\boldsymbol{h}_l = \bar{\boldsymbol{h}}_l = \frac{1}{m} \sum_{i=1}^m h_{il}$ for $l = 1, \dots, p$, \boldsymbol{X} is a matrix of order $p \times (q + 1)$ known as the design matrix or matrix of covariates, $\boldsymbol{a} = (a_0, a_1, \dots, a_q)^T$ is the vector of unknown parameters, $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_p)^T$ is the error vector with $\boldsymbol{\epsilon} \in N_p(\mathbf{0}, \sigma_{\boldsymbol{\epsilon}}^2 \boldsymbol{I}_p)$, $\sigma_{\boldsymbol{\epsilon}}^2$ is the unknown error variance parameter, and \boldsymbol{I}_p is the $p \times p$ identity matrix with ones on the main diagonal and zeros elsewhere. The variables \boldsymbol{h} and $\boldsymbol{\epsilon}$ are regarded as realizations of vectors of random variables $\mathcal{H} = (\mathcal{H}_1, \dots, \mathcal{H}_p)^T$, with $E(\mathcal{H}) = \boldsymbol{X}\boldsymbol{a}$ and $V_{\boldsymbol{h}} = V(\mathcal{H}) = \sigma_{\boldsymbol{\epsilon}}^2 \boldsymbol{I}_p$, and $\boldsymbol{\mathcal{E}} = (\mathcal{E}_1, \dots, \mathcal{E}_p)^T$, with $E(\boldsymbol{\mathcal{E}}) = \mathbf{0}$ and $V_{\boldsymbol{\epsilon}} = V(\boldsymbol{\mathcal{E}}) = \sigma_{\boldsymbol{\epsilon}}^2 \boldsymbol{I}_p$, respectively. Suppose \boldsymbol{X} :

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1q} \\ 1 & x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{p1} & x_{p2} & \cdots & x_{pq} \end{bmatrix} \quad (3)$$

is of full rank, $\text{rank}(\mathbf{X}) = q + 1$, and that $q + 1 \leq p$. Provided that $(\mathbf{X}^T \mathbf{X})^{-1}$ exists since $\mathbf{X}^T \mathbf{X}$ also has rank $q + 1 \leq p$, the ordinary least squares estimate $\boldsymbol{\alpha}$ of \mathbf{a} is:

$$\boldsymbol{\alpha} = \mathcal{A}(\boldsymbol{h}) = \mathbf{X}^\dagger \boldsymbol{h}, \quad \mathbf{X}^\dagger = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T, \quad (4)$$

which is the same as the maximum likelihood estimate of \mathbf{a} due to the assumptions of the linear model for the error term $\boldsymbol{\epsilon}$ drawn from a multivariate normal distribution with expectation $\mathbf{0}$ and covariance matrix $\sigma_\epsilon^2 \mathbf{I}_p$. Note that, $\boldsymbol{\alpha}$ is a realization of a vector of random variables $\mathbf{A} = \mathcal{A}(\boldsymbol{\mathcal{H}}) = \mathbf{X}^\dagger \boldsymbol{\mathcal{H}}$ with expectation $E(\mathbf{A}) = E(\mathbf{X}^\dagger \boldsymbol{\mathcal{H}}) = \mathbf{X}^\dagger E(\boldsymbol{\mathcal{H}}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} \mathbf{a} = \mathbf{a}$. Therefore, $\boldsymbol{\alpha}$ is an unbiased estimate of \mathbf{a} . This parameter estimate can be used to make a prediction of the mean estimate of the post-process inspection results. However, this does not provide a measure of uncertainty associated with this mean estimate.

The uncertainty of the mean estimate will depend on the uncertainty covariance of the model parameter estimate. The covariance matrix V_α of \mathbf{A} is $V_\alpha = V(\mathbf{A}) = \mathbf{X}^\dagger V(\boldsymbol{\mathcal{H}}) (\mathbf{X}^\dagger)^T = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\sigma_\epsilon^2 \mathbf{I}_p) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} = \sigma_\epsilon^2 (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} = \sigma_\epsilon^2 (\mathbf{X}^T \mathbf{X})^{-1}$. The estimate $\boldsymbol{\alpha}$ is also a sample from a multivariate normal distribution, $\boldsymbol{\alpha} \in N_{q+1}(\mathbf{a}, \sigma_\epsilon^2 (\mathbf{X}^T \mathbf{X})^{-1})$, since $\boldsymbol{\alpha}$ results from a linear transformation of the data vector \boldsymbol{h} . The maximum likelihood estimate $\hat{\sigma}_\epsilon^2 = \hat{\boldsymbol{\epsilon}}^T \hat{\boldsymbol{\epsilon}} / p$ of the error variance σ_ϵ^2 is biased, but as p increases the bias of $\hat{\sigma}_\epsilon^2$ shrinks toward 0. An unbiased estimate of the error variance σ_ϵ^2 can be obtained from the Residual Sum of Squares (RSS) divided by its degrees of freedom:

$$s_\epsilon^2 = \frac{\hat{\boldsymbol{\epsilon}}^T \hat{\boldsymbol{\epsilon}}}{p - q - 1} = \frac{(\boldsymbol{h} - \mathbf{X} \boldsymbol{\alpha})^T (\boldsymbol{h} - \mathbf{X} \boldsymbol{\alpha})}{p - q - 1}. \quad (5)$$

Therefore, the estimated covariance matrix $\hat{V}(\mathbf{A}) = s_\epsilon^2 (\mathbf{X}^T \mathbf{X})^{-1}$ of \mathbf{A} approximates the covariance matrix $V(\mathbf{A}) = \sigma_\epsilon^2 (\mathbf{X}^T \mathbf{X})^{-1}$ of \mathbf{A} . Computational methods based on matrix decompositions are used for numerical stability [49].

In this paper, a Bayesian approach is employed to first estimate the parameters \mathbf{a} and σ_{ε}^2 using only the in-process data. A Bayesian approach to statistical estimation and inference of regression models considers the parameters \mathbf{a} and σ_{ε}^2 as random variables rather than fixed, unknown quantities and hence uses probability distributions to describe information about the model parameters. These probability distributions are updated using Bayes' theorem when information from measurement data becomes available to obtain posterior distributions of the parameters. This approach can incorporate prior knowledge (before the data are observed) about the model parameters by setting up an informative prior distribution for the parameters [18]. The prior knowledge may be available from various sources of information depending on the specific problem. For example, expert knowledge and historical data can be used to obtain a prior distribution. In cases where prior knowledge may be vague or not available then, a non-informative prior distribution can be specified. The posterior distribution summarizes all the information about the model parameters after observing the data by incorporating the prior beliefs through the prior distribution and the information from the measurement data through the likelihood function. With non-informative priors, the posterior distribution is dominated by the likelihood. On the contrary, the posterior is approximately proportional to the prior when the prior contains much more precise information about the parameters than the data. The posterior distribution is proportional to the product of the prior distribution and the likelihood function and appropriately normalized to ensure that it integrates to one. The data in a regression problem include both \mathbf{h} and \mathbf{X} . Therefore, the posterior density for the unknown parameter vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}, \boldsymbol{\theta}_x)^T$ is given by [19]:

$$p(\boldsymbol{\theta}|\mathbf{h}, \mathbf{X}) \propto p(\boldsymbol{\theta})p(\mathbf{h}, \mathbf{X}|\boldsymbol{\theta}), \quad (6)$$

where $p(\boldsymbol{\theta})$ is the prior density and $p(\mathbf{h}, \mathbf{X}|\boldsymbol{\theta}) = p(\mathbf{h}|\mathbf{X}, \boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}})p(\mathbf{X}|\boldsymbol{\theta}_x)$ is the likelihood. Assuming prior independence, $p(\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}, \boldsymbol{\theta}_x) = p(\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}})p(\boldsymbol{\theta}_x)$, then, the posterior density factors as:

$$p(\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}, \boldsymbol{\theta}_x|\mathbf{h}, \mathbf{X}) = p(\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}|\mathbf{h}, \mathbf{X})p(\boldsymbol{\theta}_x|\mathbf{X}). \quad (7)$$

Since our interest lies only with $\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}} = (\mathbf{a}, \sigma_{\varepsilon}^2)$ then, the posterior density for the parameter vector $\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}$ is given by:

$$p(\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}|\mathbf{h}, \mathbf{X}) \propto p(\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}})p(\mathbf{h}|\mathbf{X}, \boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}), \quad (8)$$

where $p(\boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}})$ is the prior density and $p(\mathbf{h}|\mathbf{X}, \boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}})$ is the likelihood given by:

$$p(\mathbf{h}|\mathbf{X}, \boldsymbol{\theta}_{\mathbf{h}|\mathbf{x}}) \propto (\sigma_{\varepsilon}^2)^{-p/2} \exp \left[\frac{-(\mathbf{h} - \mathbf{X}\mathbf{a})^T(\mathbf{h} - \mathbf{X}\mathbf{a})}{2\sigma_{\varepsilon}^2} \right]. \quad (9)$$

The posterior is derived analytically when prior knowledge can be expressed with conjugate priors. Conjugate priors for the parameter vector $\theta_{h|x}$ can be given by:

$$\mathbf{a}|\sigma_{\xi}^2 \sim N_{q+1}(\boldsymbol{\alpha}_0, \sigma_{\xi}^2 \mathbf{V}_0^{-1}), \quad \sigma_{\xi}^2 \sim \text{IG}(a_0, b_0), \quad (10)$$

where the conditional prior for \mathbf{a} is a $(q + 1)$ -dimensional multivariate normal distribution with mean vector $\boldsymbol{\alpha}_0$ and covariance matrix $\sigma_{\xi}^2 \mathbf{V}_0^{-1}$ and the marginal prior for σ_{ξ}^2 is an inverse Gamma distribution with shape $a_0 > 0$ and scale $b_0 > 0$. Note, \mathbf{V}_0 is a symmetric positive definite matrix of order $(q + 1) \times (q + 1)$. A non-informative prior usually used for the normal linear regression model is the improper density $p(\mathbf{a}, \sigma_{\xi}^2) \propto 1/\sigma_{\xi}^2$.

4. Model validation using experimental multistage manufacturing data

Experimental data have been obtained from a case study concerned with multiple stages of manufacturing, such as heat treatment, grinding, hardness testing, CNC machining, and dimensional inspections [25]–[27]. The starting material blocks made of steel EN24T were heated up to 845°C using a VECSTAR furnace and then quenched in oil for hardening. After hardening, the blocks were tempered at different temperatures, including 450°C, 550°C and 650°C, to introduce variation in material properties. High temperature thermocouples were used to measure temperature gradient and temperature variation during heat treatment. Following hardening and tempering, the material blocks were grinded to improve their surface quality and surface hardness measurements of the material blocks were then obtained using a Rockwell device. A full factorial design with four factors at two levels and one center point each 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 setting the workpiece in the second orientation. During machining, an accelerometer sensor and the NI LabVIEW SignalExpress software were employed to measure tool vibration signals at a sampling rate of 10 kHz. After each experimental run, all the cutting tools used for the machining operations were inspected for wear on each flute using a Leica microscope. The cutting tools were used until they reached a certain flank wear width. The following features were generated in MATLAB from the in-process monitoring data: i) a three-state variable for the material surface hardness obtained from repeated measurements at different locations on the material block; ii) the maximum tempering temperature obtained by five K-type thermocouples; and iii) time domain features including Root-Mean-Square (RMS), sample kurtosis, sample skewness, sample variance and sample mean of vibration components V_x , V_y , and V_z . The dataset was normalized by the Euclidean norm (2-norm). After machining each side of the workpiece, OMP was performed for in-process inspection using a Renishaw OMP60 optical transmission probe. Figure 1 shows a general overview of the experimental setup and the CAD model of the product. The measurand of interest in this work is the diameter of the circle labeled in Figure 1. A Renishaw Equator 300 Extended Height Gauging System was used for post-process inspection on the shop floor. The Equator was employed using the CMM Compare method in scanning measuring mode. The CMM used to calibrate the master part was a Mitutoyo CMM located in a metrology lab.

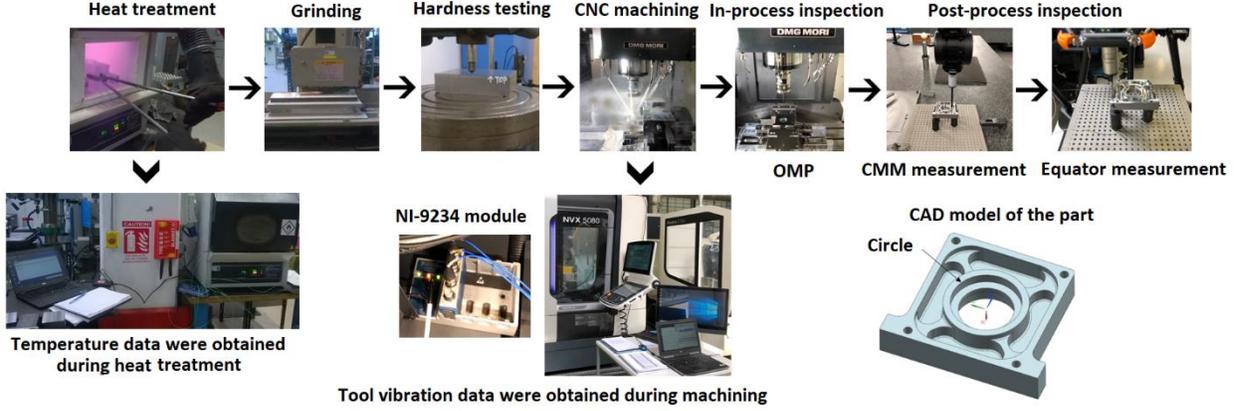


Figure 1. The experimental setup and the CAD model of the part. The subfigures indicate the different manufacturing and inspection stages with the arrows indicating the sequence of the processes.

Suppose that prior knowledge about $\theta_{h|x}$ is not available or ignored, e.g. it is considered as imprecise. Therefore, the improper prior density $p(\mathbf{a}, \sigma_{\varepsilon}^2) \propto 1/\sigma_{\varepsilon}^2$ is specified. Provided that $\mathbf{X}^T \mathbf{X}$ is invertible, with this prior distribution, the conditional posterior distribution for \mathbf{a} is a $(q + 1)$ -dimensional multivariate normal distribution:

$$\mathbf{a} | \sigma_{\varepsilon}^2, \mathbf{h}, \mathbf{X} \sim N_{q+1} \left(\boldsymbol{\alpha}, \sigma_{\varepsilon}^2 (\mathbf{X}^T \mathbf{X})^{-1} \right). \quad (11)$$

The marginal posterior distribution for σ_{ε}^2 is an inverse Gamma distribution:

$$\sigma_{\varepsilon}^2 | \mathbf{h}, \mathbf{X} \sim \text{IG} \left(\frac{p - q - 1}{2}, \frac{(\mathbf{h} - \mathbf{X}\boldsymbol{\alpha})^T (\mathbf{h} - \mathbf{X}\boldsymbol{\alpha})}{2} \right). \quad (12)$$

The marginal posterior distribution for \mathbf{a} is a $(q + 1)$ -dimensional multivariate t-distribution with $p - q - 1$ degrees of freedom:

$$\mathbf{a} | \mathbf{h}, \mathbf{X} \sim t_{q+1, p-q-1} \left(\boldsymbol{\alpha}, s_{\varepsilon}^2 (\mathbf{X}^T \mathbf{X})^{-1} \right). \quad (13)$$

Now suppose the model is applied to a new set of data $\tilde{\mathbf{X}}$ to predict unobserved data $\tilde{\mathbf{h}}$, where $\tilde{\mathbf{X}}$ is a matrix of dimension $m \times (q + 1)$ and $\tilde{\mathbf{h}} \in \tilde{\mathcal{H}}$ is a vector of dimension m . The uncertainty associated with the posterior predictive distribution for the vector of random variables $\tilde{\mathcal{H}}$ is contributed by both the model variability and finite sample size p of \mathbf{h} . Therefore, the posterior predictive distribution for $\tilde{\mathcal{H}}$ is a

m -dimensional multivariate t-distribution centered at $\tilde{\mathbf{X}}\boldsymbol{\alpha}$ with two uncertainty components and $p - q - 1$ degrees of freedom:

$$\tilde{\mathcal{H}}|\mathbf{h} \sim t_{m,p-q-1}\left(\tilde{\mathbf{X}}\boldsymbol{\alpha}, s_{\tilde{\mathcal{E}}}^2\left(\mathbf{I}_m + \tilde{\mathbf{X}}(\mathbf{X}^T\mathbf{X})^{-1}\tilde{\mathbf{X}}^T\right)\right). \quad (14)$$

The predictive model was fitted using the first four principal components extracted from the input dataset and the diameter deviations obtained from the scanning comparator measurements using a sampling point density/distance (the distance between the points on the scan path, in the current units) of 0.1 and a scanning speed of 40 mm/s. The expanded measurement uncertainties, U , for diameter for a coverage factor $k = 2$ and a confidence level of 95.45% were found to be less than 1 μm for all parts tested. The procedure followed to evaluate the uncertainty associated with this measurement process is described in [27]. Principal Component Analysis (PCA) was performed via a Singular Value Decomposition (SVD) of the input data matrix [49], [50]. Mean-centering the columns of the input data matrix was therefore a necessary pre-processing step. The leave-one-out cross-validation approach was used to evaluate the performance of the model on unseen data. The average percentage variability explained by the first four principal components during training was 96.2878% and by taking into account only the first four principal components the average reconstruction Root Mean Squared Error (RMSE) was 0.0101. The fitted predictive model was tested using the test dataset by applying the PCA transformation obtained from the training data to the test dataset. The average reconstruction RMSE considering the first four principal components obtained from the trained PCA model using the test dataset was 0.0149. Figure 2 shows the scree plot of the percentage variability explained by the first four principal components during training for a single fold data. For each validation case, sixteen parts were used for training and one part was used for testing the model and repeated across all folds. MATLAB was used to implement the PCA-based probabilistic model.

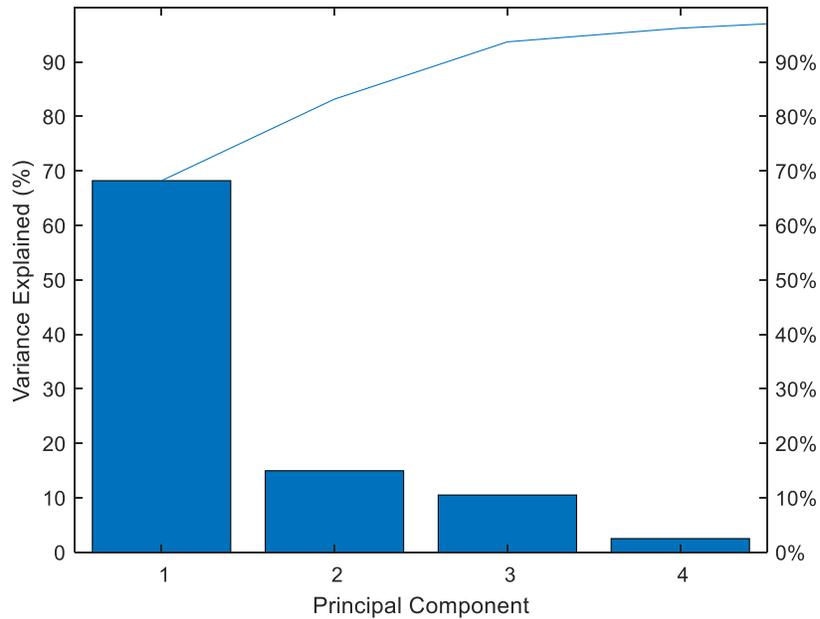


Figure 2. Scree plot for a single fold data.

Figure 3 shows the normal probability plot of the residuals of the classical linear fitted model for a single fold data. As can be seen, the model residuals follow a normal distribution with small deviations from normality. The coefficient of (multiple) determination R^2 was 0.706 and the adjusted \tilde{R}^2 was 0.599. Figure 4 shows the prior and posterior distributions of the model parameters \mathbf{a} and σ_{ϵ}^2 . Table 1 shows the Bayesian linear regression model results for a single fold data including the mean value of the parameters, the Standard Error (SE) of the parameters, and the 95% Bayesian equal-tailed Credible Interval (CI) for the parameters. Tables 2 and 3 show the cross-validation results of the PCA-based probabilistic model. The developed model was evaluated using the RMSE metric. The comparison between the training and testing data performances for PCA indicate that feature extraction process did not result in any adverse loss of information in the features from the in-process monitoring data. The cross-validation results show that the linear model can predict the diameter deviations with an average accuracy better than 6 μm . Again, the similarity of the order between the training and training errors across all parts suggest a good generalization with little overfitting to the data.

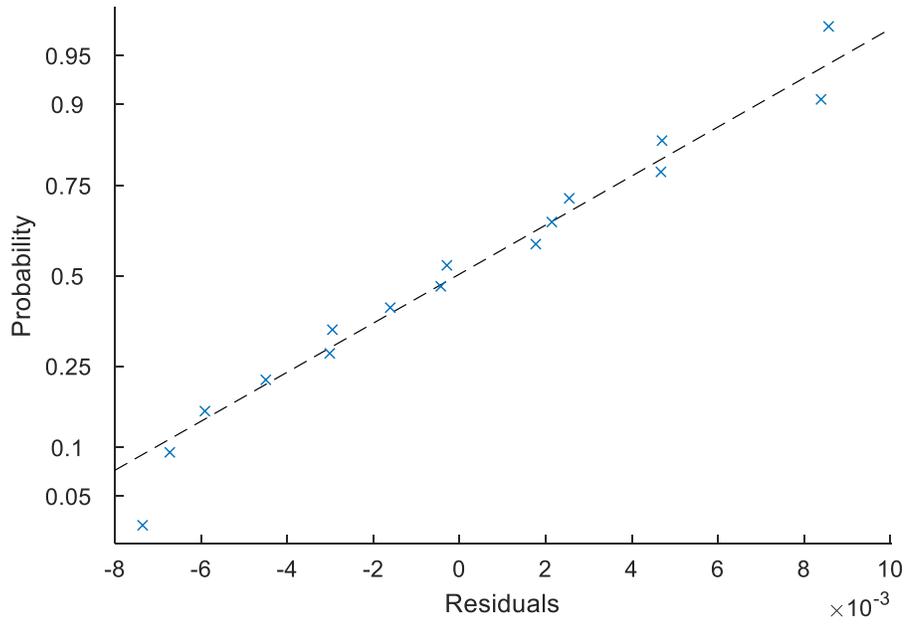


Figure 3. Normal probability plot of residuals for a single fold data.

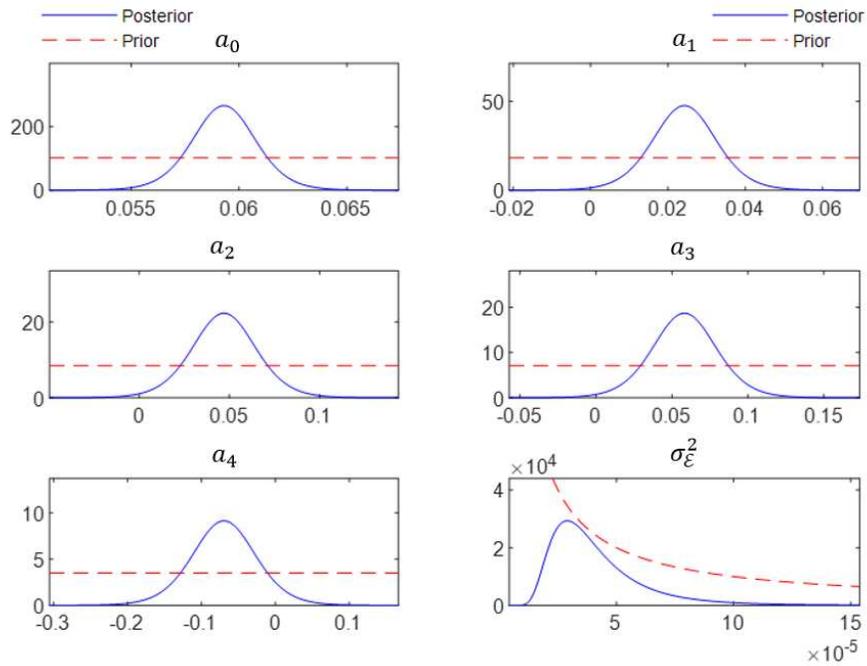


Figure 4. Prior and posterior distributions of the regression coefficients and disturbance variance for a single fold data.

Table 1. Results of the PCA-based probabilistic model for a single fold data.

	Estimate	SE	Bayesian CI95
a_0	0.05930	0.00162	[0.05608, 0.06252]
a_1	0.02427	0.00906	[0.00624, 0.04230]
a_2	0.04698	0.01934	[0.00848, 0.08548]
a_3	0.05821	0.02307	[0.01228, 0.10414]
a_4	-0.06964	0.04720	[-0.16360, 0.02433]
σ_{ξ}^2	0.00004	0.00002	[0.00002, 0.00010]

Table 2: Cross-validation results of the PCA model.

Folds	Sum of variance explained	Training reconstruction RMSE	Testing reconstruction RMSE
1	96.2111	0.0102	0.0128
2	96.2355	0.0100	0.0182
3	96.2329	0.0102	0.0120
4	96.2616	0.0101	0.0207
5	96.0725	0.0104	0.0092
6	96.2358	0.0102	0.0136
7	96.3934	0.0100	0.0174
8	96.3579	0.0099	0.0175
9	96.3920	0.0102	0.0133
10	96.0847	0.0103	0.0123
11	96.0049	0.0105	0.0072
12	96.6805	0.0095	0.0220
13	96.1143	0.0103	0.0105
14	96.3225	0.0101	0.0149
15	96.2642	0.0102	0.0121
16	96.4237	0.0100	0.0155
17	96.6048	0.0095	0.0245
Average	96.2878	0.0101	0.0149

Table 3: Cross-validation results of the predictive probabilistic model.

Folds	R^2	\tilde{R}^2	Training RMSE (mm)	Testing RMSE (mm)
1	0.706	0.599	0.0049	0.0019
2	0.664	0.542	0.0042	0.0131
3	0.746	0.654	0.0044	0.0092
4	0.736	0.639	0.0046	0.0052
5	0.711	0.605	0.0048	0.0055
6	0.740	0.646	0.0045	0.0099
7	0.695	0.584	0.0048	0.0050
8	0.778	0.697	0.0042	0.0112
9	0.726	0.627	0.0047	0.0060

10	0.711	0.605	0.0049	0.0028
11	0.692	0.580	0.0049	0.0021
12	0.707	0.600	0.0046	0.0087
13	0.700	0.591	0.0048	0.0017
14	0.743	0.650	0.0045	0.0086
15	0.710	0.605	0.0049	0.0005
16	0.708	0.602	0.0049	0.0032
17	0.592	0.443	0.0049	0.0029
Average	0.710	0.604	0.0047	0.0057

5. Combining information from different sources via the normal approximation

OMP has been used successfully in a variety of machining processes to replace hard gauging and improve the productivity and capability of the machining process, provided that the machine performs with sufficient accuracy and repeatability. It can provide feedback to the machining process to compensate for inherent variations, such as tool wear, and verify the process before the machined part is moved. However, it provides measurement results with high levels of uncertainties, which can compromise the reliability of the inspection process. Depending on the accuracy requirements of an application, independent measurements, such as CMM measurements, may be required to supplement this inspection strategy.

Consider now a Bayesian information fusion approach to update the probability distribution $p(\tilde{\mathbf{h}}|\mathbf{h})$ of the end product quality by incorporating new information from subsequent measurements, such as OMP, $\mathbf{h}_{OMP} \in \mathcal{H}$, as the product is manufactured. A Renishaw OMP60 optical transmission probe with tungsten carbide stylus (50 mm overall length) and ruby ball (2 mm ball diameter) was used for OMP. The unidirectional repeatability of the OMP60 probe is $1 \mu\text{m} 2\sigma$. A circle was fitted to the OMP data in the xy -plane in MATLAB in order to evaluate the diameter of the circular feature of interest from the OMP coordinate data. Given a two-dimensional set of m data points $\mathcal{P} = \{\mathbf{p}_i = (x_i, y_i)^T\}_1^m$, with $m \geq n$, representing a circle specified by parameters $\mathbf{b} = (x_0, y_0, r_0)^T$, where (x_0, y_0) is its centre coordinates, r_0 is its radius, and n is the number of parameters and thus in this case $n = 3$, the least squares best-fit circle is found by solving the optimization problem:

$$\min_{\mathbf{b}} \sum_{i=1}^m d^2(\mathbf{p}_i, \mathbf{b}), \quad (15)$$

where $d(\mathbf{p}_i, \mathbf{b}) = d_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - r_0$ is the (signed) distance from a point \mathbf{p}_i to the circle specified by \mathbf{b} , with $d_i > 0$ for points outside the circle and $d_i < 0$ for points inside it. One of the most common approaches in solving this optimization problem is the Gauss-Newton algorithm [51]. This

algorithm requires the $m \times n$ Jacobian matrix \mathbf{J} of partial derivatives of d_i with respect to the parameters \mathbf{b} [52], [53]:

$$\begin{aligned}\frac{\partial d_i}{\partial x_0} &= -\frac{(x_i - x_0)}{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}} \\ \frac{\partial d_i}{\partial y_0} &= -\frac{(y_i - y_0)}{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}} \\ \frac{\partial d_i}{\partial r_0} &= -1.\end{aligned}\tag{16}$$

Note, for each part/circle, eight points were taken as such a sample size of probing points can provide sufficient information on diameter deviation and is practical in order to maintain inspection cycle time within desired limits. Let h_{OMP} be a single scalar circular feature characteristic (diameter) obtained from an OMP inspection cycle with $h_{OMP} \in N(\mu_{\mathcal{H}}, \sigma_{\mathcal{H}}^2)$. If the variance $\sigma_{\mathcal{H}}^2$ is a known constant then, the likelihood is represented by:

$$p(h_{OMP}|\mu_{\mathcal{H}}) = \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{H}}} \exp\left[\frac{-1}{2\sigma_{\mathcal{H}}^2}(h_{OMP} - \mu_{\mathcal{H}})^2\right].\tag{17}$$

The likelihood, $p(h_{OMP}|\mu_{\mathcal{H}})$, considered as a function of the unknown mean parameter $\mu_{\mathcal{H}} = E(\mathcal{H})$ of the measurand \mathcal{H} , is a one-dimensional exponential family and thus, a conjugate prior density for this likelihood is the normal density, i.e., $\mu_{\mathcal{H}} \sim N(\mu_0, \sigma_0^2)$. For the fusion process the posterior predictive distribution for $\tilde{\mathcal{H}}$, $p(\tilde{h}|\mathbf{h})$, becomes the prior density for $\mu_{\mathcal{H}}$, $p(\mu_{\mathcal{H}})$, and can be well approximated by a normal density as the degrees of freedom of the t-distribution increase. Thus, the prior for $\mu_{\mathcal{H}}$ is chosen as:

$$p(\mu_{\mathcal{H}}) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left[\frac{-1}{2\sigma_0^2}(\mu_{\mathcal{H}} - \mu_0)^2\right],\tag{18}$$

where $\mu_0 = \tilde{\mathbf{X}}\boldsymbol{\alpha}$ and $\sigma_0^2 = s_{\tilde{\boldsymbol{\varepsilon}}}^2 (\mathbf{I}_m + \tilde{\mathbf{X}}(\mathbf{X}^T\mathbf{X})^{-1}\tilde{\mathbf{X}}^T)$. The influence of this approximation on the posterior inferences decreases as the sample size p increases. However, such an approximation may not be accurate for small sample sizes. With a conjugate prior density, the posterior density for $\mu_{\mathcal{H}}$ is also a normal density $\mu_{\mathcal{H}}|h_{OMP}, \tilde{\mathbf{X}} \sim N(\mu_1, \sigma_1^2)$.

The Bayesian inference results in the posterior mean μ_1 being a precision-weighted average of the prior mean μ_0 and the observed value \bar{h}_{OMP} , and the posterior precision (inverse of the variance) σ_1^{-2} being the sum of the prior precision σ_0^{-2} and the observation precision $\sigma_{\mathcal{H}}^{-2}$:

$$\mu_1 = w\mu_0 + (1 - w)\bar{h}_{OMP}, \quad \sigma_1^{-2} = \sigma_0^{-2} + \sigma_{\mathcal{H}}^{-2}, \quad (19)$$

where $w = \sigma_0^{-2}/(\sigma_0^{-2} + \sigma_{\mathcal{H}}^{-2}) \in (0,1)$. If the product is measured on the machine m_{OMP} times independently, under repeatability conditions then, the joint likelihood is the product of the individual likelihoods:

$$\begin{aligned} p(\bar{h}_{1OMP}, \dots, \bar{h}_{m_{OMP}} | \mu_{\mathcal{H}}) &= p(\bar{h}_{1OMP} | \mu_{\mathcal{H}}) \times \dots \times p(\bar{h}_{m_{OMP}} | \mu_{\mathcal{H}}) \\ &= \prod_{i=1}^{m_{OMP}} \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{H}}} \exp\left[\frac{-1}{2\sigma_{\mathcal{H}}^2} (\bar{h}_{iOMP} - \mu_{\mathcal{H}})^2\right] \propto \exp\left[\frac{-1}{2\sigma_{\mathcal{H}}^2} \sum_{i=1}^{m_{OMP}} (\bar{h}_{iOMP} - \mu_{\mathcal{H}})^2\right]. \end{aligned} \quad (20)$$

The new measurements based on OMP are subject to both random and systematic effects and are considered as the likelihood function that is also Gaussian. Therefore, $\mu_{\mathcal{H}}$ has posterior $\mu_{\mathcal{H}} | \bar{h}_{OMP}, \bar{X} \sim N(\mu_{m_{OMP}}, \sigma_{m_{OMP}}^2)$ with mean $\mu_{m_{OMP}}$ and precision $\sigma_{m_{OMP}}^{-2}$:

$$\mu_{m_{OMP}} = \frac{m_{OMP}\sigma_{\mathcal{H}}^{-2}\bar{h}_{OMP} + \sigma_0^{-2}\mu_0}{m_{OMP}\sigma_{\mathcal{H}}^{-2} + \sigma_0^{-2}}, \quad \sigma_{m_{OMP}}^{-2} = m_{OMP}\sigma_{\mathcal{H}}^{-2} + \sigma_0^{-2}, \quad (21)$$

where $\bar{h}_{OMP} = \frac{1}{m_{OMP}} \sum_{i=1}^{m_{OMP}} \bar{h}_{iOMP}$. The maximum likelihood estimate of $\mu_{\mathcal{H}}$ is simply the sample mean \bar{h}_{OMP} .

The measurement uncertainty associated with OMP was calculated based on the uncertainty evaluation methodology given in ISO 15530-3 [54], which provides an experimental technique for CMM measurement uncertainty evaluation with calibrated parts of similar dimension and geometry. Measurement uncertainty and measurement error are often mistakenly used interchangeably because they are thought to be synonymous. However, this is incorrect since measurement error is the measurement result minus the true value of the measurand while measurement uncertainty is an attribute of the measurement result, which quantifies the doubt about its validity [55]. Three uncertainty contributors were considered to evaluate the uncertainty associated with OMP: i) the standard uncertainty, $u(cal)$, associated with the uncertainty of the CMM calibration of the master part; ii) the standard uncertainty, $u(p)$, associated with the OMP measurement procedure; and iii) the standard uncertainty, $u(b)$, associated with the systematic error component, $b = |\bar{y} - y_{cal}|$, of the OMP

measurement process evaluated using the master part, where \bar{y} denotes the mean of the measured values obtained from OMP and ψ_{cal} is the CMM calibrated value of the same part and measurand. The uncertainty component $u(cal)$ brought-in from the calibration of the master part was evaluated by the experimental standard deviation of the mean, $s(\bar{y}_{cal}) = s_{\psi_{cal}}/\sqrt{m_{cal}}$, using the CMM measurements on the master part, where $s_{\psi_{cal}} = \sqrt{\frac{1}{m_{cal}-1} \sum_{i=1}^{m_{cal}} (\psi_{i_{cal}} - \bar{y}_{cal})^2}$. The uncertainty components, $u(p)$ and $u(b)$, were calculated by the experimental standard deviation of the mean using the OMP measurements on the test part and master part, respectively. The uncertainty associated with OMP, $u(\psi)$, for a coverage probability of 68.27% can be given by:

$$u(\psi) = \sqrt{u^2(cal) + u^2(p) + u^2(b)} + b. \quad (22)$$

Table 4 shows the residual values between the Equator (post-process inspection results) and the PCA-based probabilistic model, the Equator and the OMP, and the Equator and the Bayesian information fusion algorithm, obtained from the cross-validation process. MATLAB was used to implement the proposed algorithm. Figure 5 shows the prior distribution, likelihood, and posterior distribution of the end product quality for a single fold data. The resulting posterior is a normal distribution representing a compromise between the new information obtained from OMP and machine learning-based prediction information. Conjugacy is an important concept in Bayesian data analysis [56]. If more information is available at a later manufacturing stage or time, the present posterior distribution can be used as the new prior distribution and the new information can be considered as the new likelihood function. In other words, the posterior distribution obtained at manufacturing step i acts as a prior distribution at step $i + 1$. The sequential nature of Bayesian approach allows us to quantify many sources of information and uncertainties in the form of posterior distributions. The fusion is made possible by the assumption of independence in the measurement process from which the probability distributions are obtained. Our $100(1 - \mathcal{Q})\%$ posterior interval for $\mu_{\mathcal{H}}$ is $\mu_{m_{OMP}} \pm z \frac{\mathcal{Q}}{2} \times \sigma_{m_{OMP}}$, where the z -value can be found in the standard normal distribution table and \mathcal{Q} defines the amount of uncertainty and is predetermined. If an estimate $s_{\mathcal{H}}^2 = \frac{1}{m_{OMP}-1} \sum_{i=1}^{m_{OMP}} (\mathcal{H}_{i_{OMP}} - \bar{\mathcal{H}}_{OMP})^2$ of $\sigma_{\mathcal{H}}^2$ is used then, the correct posterior interval for $\mu_{\mathcal{H}}$ is $\mu_{m_{OMP}} \pm t \frac{\mathcal{Q}}{2} \times \sigma_{m_{OMP}}$, where similarly the t -value can be found in the Student's t -distribution table. Although linear models do not have a high precision of predicting the mean value of the health parameter, this can be improved with a nonlinear model. The key concept of the proposed method was to introduce the inclusion of the uncertainty in the model prediction and how this can be harnessed in improving the estimation of the health parameters with additional information. The expression for the uncertainty measure for a linear model can be computed efficiently and hence its choice in this research work. As can be seen from Table 4, the fused results show good performance in health parameter estimation. The RMSE performance metric of the linear model (prior prediction) is 6.7847 μm and that of OMP (likelihood) is 16.9434 μm , which is significantly higher compared to the linear model. OMP using TTPs in the machine spindle is susceptible to a wide range of uncertainty sources and the main problem with this is that the measurement process is not independent of the machining process. The RMSE of the information fusion algorithm (posterior prediction) is 6.2875 μm , which is lower than both of that of the linear model and OMP, thus demonstrating the value of fusing information from in-process monitoring and inspection data.

Table 4: Residuals of the different methods used for obtaining the product health parameters.

Folds	PCA-based probabilistic model and Equator (μm)	OMP and Equator (μm)	Information fusion algorithm and Equator (μm)
1	1.8719	12.4367	3.5299
2	13.0625	9.7696	12.6638
3	9.2300	46.2653	1.7815
4	5.1790	27.8882	8.4883
5	5.5121	0.8522	4.5178
6	9.8651	3.5189	8.9802
7	4.9658	5.0737	4.9825
8	11.2118	21.2005	12.4572
9	5.9755	21.1973	1.8981
10	2.8374	15.9113	4.8916
11	2.1355	7.1975	0.6584
12	8.6556	3.2939	7.8811
13	1.6696	6.2824	2.3780
14	8.5706	17.4240	4.9172
15	0.4836	7.3260	0.7457
16	3.2275	2.1101	3.0508
17	2.8641	7.5718	3.6269
Average	5.7246	12.6658	5.1441
RMSE	6.7847	16.9434	6.2875

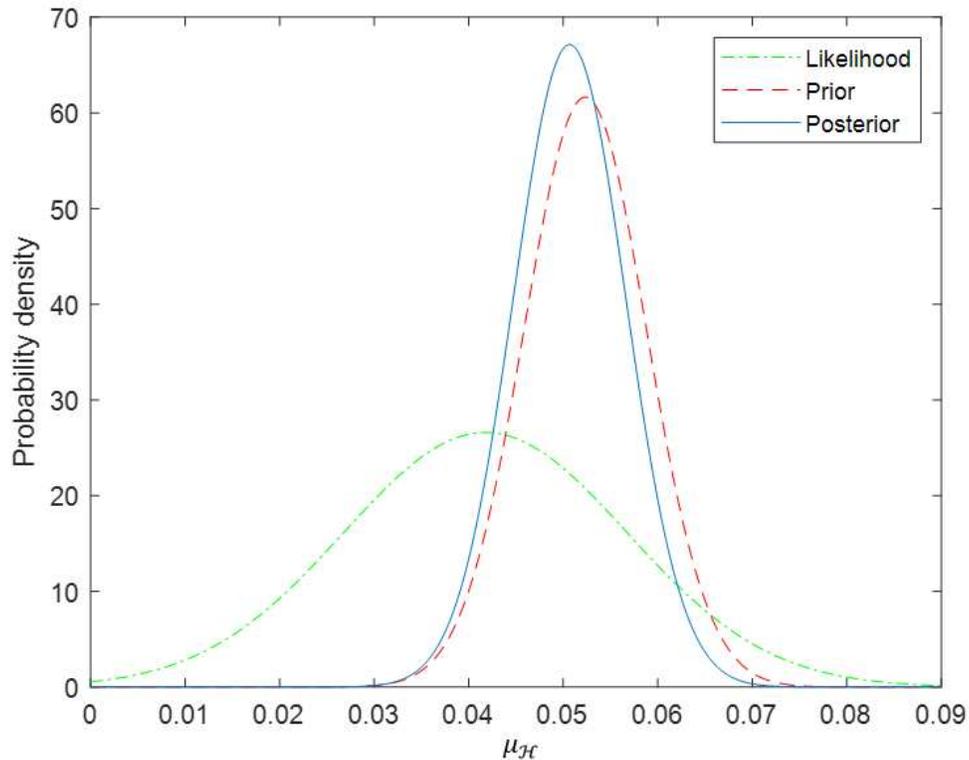


Figure 5. Prior, likelihood, and posterior distribution for a single fold data.

6. Discussion and conclusions

Manufacturing processes usually involve a series of operations fraught with several sources of errors. Incorporating intelligence in manufacturing systems can help increase production efficiency and reduce material waste, environmental impacts, and human intervention in manufacturing operations. Existing product health monitoring systems are usually limited to monitoring only the machining process to identify the end product quality and do not update their predictions when new metrological information, such as On-Machine Measurements (OMMs), becomes available at a later manufacturing stage. This paper has been concerned with developing a Bayesian information fusion approach to update the probability distribution of the end product quality given new information from On-Machine Probing (OMP) inspection data. Bayesian methods offer a natural framework to model uncertainty and combine information extracted from different types of data sources. The uncertainty associated with OMP was calculated following the uncertainty evaluation methodology given in ISO 15530-3. Prior knowledge about the product condition was obtained from a linear probabilistic model, but other machine learning models, such as Artificial Neural Networks (ANNs), can be developed to predict the dimensional metrology characteristics of interest. The posterior distribution marrying the new information obtained from OMP with machine learning-based prediction information clearly represents a compromise between the two sources of information.

The proposed methods have some limitations. The use of a linear model to capture the relationship between post-process inspection and in-process monitoring data can be improved upon by a nonlinear model that has the flexibility of representing more complex relationships [57]. The approximation of the probability distributions as Gaussians is also somewhat limiting. If extensive data are available, then these distributions can be more accurately represented by nonparametric methods. This in turn will require more computationally involved Bayesian inference algorithms, such as the use of sampling methods to compute the posterior probability distributions [18], [19]. Finally, the choice of the health parameter estimate may also require a robust estimate from the non-Gaussian posterior distribution.

The Bayesian fusion method proposed to predict an improved health parameter estimate introduces a novel means to include in-process inspection data with the in-process monitoring data. The approach has been demonstrated through a case study involving multiple manufacturing stages. The methods proposed here are generic to other processes, such as in the case of additional sensors included for inspection of the products as distinct from the process monitoring system. Industry 4.0 is driving this transformation in which more sensors are added, and intelligent decisions are made to increase throughput and avoid highly costly and time-consuming accurate inspection processes taking place in controlled environments. Future work will look to apply the proposed method to laser welding and additive manufacturing processes.

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] Y. Altintas and A. A. Ber, "Manufacturing automation: metal cutting mechanics, machine tool vibrations, and CNC design," *Appl. Mech. Rev.*, vol. 54, no. 5, pp. B84–B84, 2001.
- [2] T. L. Schmitz and K. S. Smith, *Machining dynamics*. Springer, 2014.
- [3] R. Ramesh, M. A. Mannan, and A. N. Poo, "Error compensation in machine tools—a review: part I: geometric, cutting-force induced and fixture-dependent errors," *Int. J. Mach. Tools Manuf.*, vol. 40, no. 9, pp. 1235–1256, 2000.
- [4] F. A. Niaki, M. Michel, and L. Mears, "State of health monitoring in machining: Extended Kalman filter for tool wear assessment in turning of IN718 hard-to-machine alloy," *J. Manuf. Process.*, vol. 24, pp. 361–369, 2016.
- [5] B. Lu and X. Zhou, "Quality and reliability oriented maintenance for multistage manufacturing systems subject to condition monitoring," *J. Manuf. Syst.*, vol. 52, pp. 76–85, 2019.
- [6] J.-P. Loose, S. Zhou, and D. Ceglarek, "Kinematic analysis of dimensional variation propagation for multistage machining processes with general fixture layouts," *IEEE Trans. Autom. Sci. Eng.*, vol. 4, no. 2, pp. 141–152, 2007.
- [7] J.-P. Loose, Q. Zhou, S. Zhou, and D. Ceglarek, "Integrating GD&T into dimensional variation models for multistage machining processes," *Int. J. Prod. Res.*, vol. 48, no. 11, pp. 3129–3149, 2010.

- [8] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," *Manuf. Lett.*, vol. 3, pp. 18–23, 2015.
- [9] S. Wang, J. Wan, D. Zhang, D. Li, and C. Zhang, "Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination," *Comput. networks*, vol. 101, pp. 158–168, 2016.
- [10] L. Lv, Z. Deng, T. Liu, Z. Li, and W. Liu, "Intelligent technology in grinding process driven by data: A review," *J. Manuf. Process.*, vol. 58, pp. 1039–1051, 2020.
- [11] A. Weckenmann *et al.*, "Multisensor data fusion in dimensional metrology," *CIRP Ann.*, vol. 58, no. 2, pp. 701–721, 2009.
- [12] B. M. Colosimo, M. Pacella, and N. Senin, "Multisensor data fusion via Gaussian process models for dimensional and geometric verification," *Precis. Eng.*, vol. 40, pp. 199–213, 2015.
- [13] M. Liggins II, D. Hall, and J. Llinas, *Handbook of multisensor data fusion: theory and practice*. CRC press, 2017.
- [14] J. V. Abellan-Nebot and F. R. Subirón, "A review of machining monitoring systems based on artificial intelligence process models," *Int. J. Adv. Manuf. Technol.*, vol. 47, no. 1, pp. 237–257, 2010.
- [15] P. Oborski, "Developments in integration of advanced monitoring systems," *Int. J. Adv. Manuf. Technol.*, vol. 75, no. 9–12, pp. 1613–1632, 2014.
- [16] B. M. Colosimo and E. Del Castillo, *Bayesian process monitoring, control and optimization*. CRC Press, 2006.
- [17] M. Papananias, T. E. McLeay, M. Mahfouf, and V. Kadirkamanathan, "A Bayesian framework to estimate part quality and associated uncertainties in multistage manufacturing," *Comput. Ind.*, vol. 105, 2019, doi: 10.1016/j.compind.2018.10.008.
- [18] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, 3rd ed. CRC Press, 2014.
- [19] S. Jackman, *Bayesian analysis for the social sciences*, vol. 846. John Wiley & Sons, 2009.
- [20] H. S. Migon, D. Gamerman, and F. Louzada, *Statistical inference: an integrated approach*. CRC press, 2014.
- [21] J. P. Choi, B. K. Min, and S. J. Lee, "Reduction of machining errors of a three-axis machine tool by on-machine measurement and error compensation system," *J. Mater. Process. Technol.*, vol. 155, pp. 2056–2064, 2004.
- [22] A. B. Forbes, A. Mengot, and K. Jonas, "Uncertainty associated with coordinate measurement in comparator mode," *Laser Metrol. Mach. Perform. XI, LAMDAMAP*, pp. 150–159, 2015.
- [23] R. J. Hocken and P. H. Pereira, *Coordinate measuring machines and systems*, vol. 6. CRC press Boca Raton, FL, 2012.
- [24] A. Forbes, M. Papananias, A. Longstaff, S. Fletcher, A. Mengot, and K. Jonas, "Developments in automated flexible gauging and the uncertainty associated with comparative coordinate measurement," 2016.
- [25] M. Papananias, T. E. McLeay, M. Mahfouf, and V. Kadirkamanathan, "An intelligent metrology

- informatics system based on neural networks for multistage manufacturing processes,” in *Procedia CIRP*, 2019, vol. 82, doi: 10.1016/j.procir.2019.04.148.
- [26] M. Papananias, O. Obajemu, T. E. McLeay, M. Mahfouf, and V. Kadirkamanathan, “Development of a new machine learning-based informatics system for product health monitoring,” in *Procedia CIRP*, 2020, vol. 93, doi: 10.1016/j.procir.2020.03.075.
- [27] M. Papananias, T. E. McLeay, O. Obajemu, M. Mahfouf, and V. Kadirkamanathan, “Inspection by exception: A new machine learning-based approach for multistage manufacturing,” *Appl. Soft Comput. J.*, vol. 97, 2020, doi: 10.1016/j.asoc.2020.106787.
- [28] M. Papananias, S. Fletcher, A. P. Longstaff, and A. B. Forbes, “Uncertainty evaluation associated with versatile automated gauging influenced by process variations through design of experiments approach,” *Precis. Eng.*, vol. 49, 2017, doi: 10.1016/j.precisioneng.2017.04.007.
- [29] T. Özel and Y. Karpat, “Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks,” *Int. J. Mach. tools Manuf.*, vol. 45, no. 4–5, pp. 467–479, 2005.
- [30] E. G. Plaza, P. J. N. López, and E. M. B. González, “Efficiency of vibration signal feature extraction for surface finish monitoring in CNC machining,” *J. Manuf. Process.*, vol. 44, pp. 145–157, 2019.
- [31] D. R. Salgado, F. J. Alonso, I. Cambero, and A. Marcelo, “In-process surface roughness prediction system using cutting vibrations in turning,” *Int. J. Adv. Manuf. Technol.*, vol. 43, no. 1, pp. 40–51, 2009.
- [32] P. B. Huang, “An intelligent neural-fuzzy model for an in-process surface roughness monitoring system in end milling operations,” *J. Intell. Manuf.*, vol. 27, no. 3, pp. 689–700, 2016.
- [33] P. Kovac, D. Rodic, V. Pucovsky, B. Savkovic, and M. Gostimirovic, “Application of fuzzy logic and regression analysis for modeling surface roughness in face milling,” *J. Intell. Manuf.*, vol. 24, no. 4, pp. 755–762, 2013.
- [34] G. Bolar, A. Das, and S. N. Joshi, “Measurement and analysis of cutting force and product surface quality during end-milling of thin-wall components,” *Measurement*, vol. 121, pp. 190–204, 2018.
- [35] C. Han, M. Luo, and D. Zhang, “Optimization of varying-parameter drilling for multi-hole parts using metaheuristic algorithm coupled with self-adaptive penalty method,” *Appl. Soft Comput.*, vol. 95, p. 106489, 2020.
- [36] J. Moore, J. Stammers, and J. Dominguez-Caballero, “The application of machine learning to sensor signals for machine tool and process health assessment,” *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, vol. 235, no. 10, pp. 1543–1557, 2021.
- [37] M. Correa, C. Bielza, and J. Pamies-Teixeira, “Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process,” *Expert Syst. Appl.*, vol. 36, no. 3, pp. 7270–7279, 2009.
- [38] O. F. Beyca, P. K. Rao, Z. Kong, S. T. S. Bukkapatnam, and R. Komanduri, “Heterogeneous sensor data fusion approach for real-time monitoring in ultraprecision machining (UPM) process using non-parametric Bayesian clustering and evidence theory,” *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 2, pp. 1033–1044, 2015.

- [39] J. Karandikar, A. Honeycutt, T. Schmitz, and S. Smith, "Stability boundary and optimal operating parameter identification in milling using Bayesian learning," *J. Manuf. Process.*, vol. 56, pp. 1252–1262, 2020.
- [40] N. Wang, G. Zhang, L. Ren, W. Pang, and Y. Wang, "Vision and sound fusion-based material removal rate monitoring for abrasive belt grinding using improved LightGBM algorithm," *J. Manuf. Process.*, vol. 66, pp. 281–292, 2021.
- [41] Q. Nazir and C. Shao, "Online tool condition monitoring for ultrasonic metal welding via sensor fusion and machine learning," *J. Manuf. Process.*, vol. 62, pp. 806–816, 2021.
- [42] M. A. Atoui, S. Verron, and A. Kobi, "A Bayesian network dealing with measurements and residuals for system monitoring," *Trans. Inst. Meas. Control*, vol. 38, no. 4, pp. 373–384, 2016.
- [43] S. Zhao, Y. Ma, and B. Huang, "Probabilistic monitoring of sensors in state-space with variational Bayesian inference," *IEEE Trans. Ind. Electron.*, vol. 66, no. 3, pp. 2154–2163, 2018.
- [44] S. Du, X. Yao, and D. Huang, "Engineering model-based Bayesian monitoring of ramp-up phase of multistage manufacturing process," *Int. J. Prod. Res.*, vol. 53, no. 15, pp. 4594–4613, 2015.
- [45] S. Riaz, M. Riaz, Z. Hussain, and T. Abbas, "Monitoring the performance of Bayesian EWMA control chart using loss functions," *Comput. Ind. Eng.*, vol. 112, pp. 426–436, 2017.
- [46] K. Ghosh, Y. S. Ng, and R. Srinivasan, "Evaluation of decision fusion strategies for effective collaboration among heterogeneous fault diagnostic methods," *Comput. Chem. Eng.*, vol. 35, no. 2, pp. 342–355, 2011.
- [47] F. Zhang and Z. Ge, "Decision fusion systems for fault detection and identification in industrial processes," *J. Process Control*, vol. 31, pp. 45–54, 2015.
- [48] R. M. Barker, M. G. Cox, A. B. Forbes, and P. M. Harris, "Software Support for Metrology Best Practice Guide No. 4. Discrete modelling and experimental data analysis," 2007.
- [49] G. H. Golub and C. F. Van Loan, "Matrix computations 4th edition the johns hopkins university press," *Balt. MD*, 2013.
- [50] I. Jolliffe and M. Lovric, "International encyclopedia of statistical science," in *Principal Component Analysis*, Springer Berlin Heidelberg Berlin, Heidelberg, 2011, pp. 1094–1096.
- [51] P. E. Gill, W. Murray, and M. H. Wright, *Practical optimization*. SIAM, 2019.
- [52] A. B. Forbes, "Uncertainty evaluation associated with fitting geometric surfaces to coordinate data," *Metrologia*, vol. 43, no. 4, p. S282, 2006.
- [53] A. B. Forbes, "Least-Square Best-Fit geometric elements. 1989," *Natl. Phys. Lab. Teddington, United Kingdom*.
- [54] 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."
- [55] "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."

- [56] W. M. Bolstad, *Understanding computational Bayesian statistics*, vol. 644. John Wiley & Sons, 2009.
- [57] C. M. Bishop, "Pattern recognition," *Mach. Learn.*, vol. 128, no. 9, 2006.