



# Quality monitoring in multistage manufacturing systems by using machine learning techniques

Mohamed Ismail<sup>1</sup> · Noha A. Mostafa<sup>2,3</sup> · Ahmed El-assal<sup>1</sup>

Received: 2 March 2021 / Accepted: 22 May 2021 / Published online: 2 July 2021  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

Manufacturing and production processes have become more complicated and usually consist of multiple stages to meet customers' requirements. This poses big challenges for quality monitoring due to the vast amount of data and the interactive effects of many factors on the final product quality. This research introduces a smart real-time quality monitoring and inspection framework capable of predicting and determining the quality deviations for complex and multistage manufacturing systems as early as possible; introduces a hybrid quality inspection approach based on both predictive models and physical inspection in order to enhance the quality monitoring process, save resources, reduce inspection time and costs. Several supervised and unsupervised machine learning techniques such as support vector machine, random forest, artificial neural network, principal component analysis were used to build the quality monitoring model with considering the cumulative effects of different manufacturing stages and the unbalance and dynamic nature of the manufacturing processes. A complex semiconductor manufacturing dataset was used to verify and assess the performance of the proposed framework. The results prove the ability of the suggested framework to enhance the quality monitoring process in multistage manufacturing systems and the ability of the hybrid quality inspection approach to reduce the inspection volume and cost.

**Keywords** Multistage manufacturing · Quality prediction · Quality monitoring · Industry 4.0 · Machine learning · Smart manufacturing

## Introduction

Product quality is a crucial element for manufacturing organizations to appraise their production capability and enhance their market competitiveness. Traditional statistical process control (SPC) methods such as control charts have been widely used to detect defects due to their applicability and simplicity. The successful implementation of traditional SPC

methods is associated with low complexity and stable manufacturing processes (Bai et al., 2019). Manufacturing and production processes have become more complicated with existing of high dimension variables, dynamic and uncertain environments (Cheng et al., 2018; Ou et al., 2020). Manufacturing and production processes usually consist of multiple manufacturing stages to produce complex products such as semiconductor manufacturing and automobile assembling. In a multistage manufacturing system (MMS), Many factors (e.g., equipment and manufacturing variables) have interactive and accumulative effects on final product quality (Kao et al., 2017). Consequently, traditional SPC techniques become insufficient in dealing with current problems and challenges (Kim et al., 2012; Wuest et al., 2014).

The development of data mining and analytics techniques offers a promising solution to enhance the quality monitoring processes and the industrial environment. Machine learning (ML) works as a computational engine for hidden pattern recognition and data mining as it can deal with complex and high-dimensional data (Ge et al., 2017; Tao et al., 2018). ML techniques can transform the vast amount of data

✉ Noha A. Mostafa  
noha.mostafa@bue.edu.eg; namostafa@eng.zu.edu.eg

Mohamed Ismail  
Mohamed.allam@bhit.bu.edu.eg

Ahmed El-assal  
Ahmed.El-Assal@bhit.bu.edu.eg

<sup>1</sup> Department of Mechanical Engineering, Benha University, Benha, Qalubia 13518, Egypt

<sup>2</sup> Department of Mechanical Engineering, The British University in Egypt, Shorouk City, Cairo 11837, Egypt

<sup>3</sup> Department of Industrial Engineering, Zagazig University, Zagazig, Sharkia 44519, Egypt

gathered by smart sensors and other data collection techniques throughout all the manufacturing stages into useful information that explain the uncertainties, analyze the complex and hidden relations and help in making more informed actions and decisions related to quality (Cohen & Singer, 2021; Lee et al., 2013). Henceforth, ML can be seen as a key enabler of Industry 4.0 and what is called ‘Smart factories’ that make use of the advances in information and communication technologies. In a recent work by Diez-Olivan et al. (2019), the potential of using ML in the context of Industry 4.0 was expanded; quality inspection and process monitoring were among the main issues that they have addressed. As the product quality depends on how it was produced, ML can build a predictive model maps between the manufacturing variables and the quality characteristics or the quality state of produced products (Lieber et al., 2013). The output from this model is a real-time quality index, which indicates the manufacturing process’s current performance so that the defects and potential failures can be detected and avoided (Kao et al., 2017).

There are seven major challenges that are faced when building a quality monitoring model for multistage manufacturing systems, they can be summarized as follows:

- the curse of dimensionality due to the vast amount of data gathered by smart sensors along the manufacturing chain (Cheng et al., 2018);
- the accumulative and interactive effect of the different workstations on the quality of the final product quality;
- the insufficient instances or samples of minority (defective) class samples, this happens as the failure instances do not happen frequently;
- some values may be missing due to poor calibration and errors in measurement (Lee et al., 2019);
- issues with the random sampling of products from the lot according to the inspection protocol;
- the changes in data characteristics, and the appearance of new patterns;
- the difficulties of technical implementation of the predictive model in quality monitoring and inspection (Schmitt et al., 2020).

Most of the existing quality monitoring models concentrate on a single manufacturing stage and do not estimate the product quality until the end of the manufacturing sequence (Arif et al., 2013a, b). Waiting until the end of the manufacturing chain can cause a waste of resources and time; ignoring the relationships between manufacturing stages can negatively affect the performance of the quality monitoring models.

This paper aims to introduce a smart real-time quality monitoring and inspection system that uses different machine learning techniques to handle complex and multistage

manufacturing processes with considering the accumulative effects of workstations and the unbalanced and dynamic nature of the manufacturing processes. Also, a hybrid quality inspection strategy based on both physical inspection and predictive model was introduced to ensure high and constant product quality and reduce the inspection cost.

The remainder of this paper is structured as follows: in Sect. 2, a brief literature review is given on state-of-the-art works of quality monitoring using machine learning techniques and the major challenges that affect building an effective quality monitoring model in MMS; In Sect. 3, the proposed quality monitoring framework is presented; In Sect. 4, a case study is presented to assess the performance of the proposed framework, and in Sect. 5, the conclusion and future work are presented.

## Literature review

### Machine learning

Machine Learning (ML) is a subfield of artificial intelligence (AI) that enables engineers to analyze and interpret the vast amount of data generated due to the development of IoT technology and gathered by smart sensors and other data collection techniques. ML techniques play a vital role in information extraction, pattern recognition and data analysis to make more informed actions and decisions (Lee et al., 2013). ML has successfully applied in many industrial fields and applications such as predictive maintenance, fault detection and diagnosis, quality monitoring, prediction and improvement and manufacturing process optimization (Aydemir & Acar, 2020; Escobar et al., 2021; Ge et al., 2017; Köksal et al., 2011; Lee et al., 2020, 2021; Nti et al., 2021; Quatrini et al., 2020; Weichert et al., 2019; Yin et al., 2018; Zhang et al., 2019). ML can be divided into unsupervised, supervised, semi-supervised, reinforcement and active learning (Ge et al., 2017).

Unsupervised learning is applied when the dataset contains only training samples  $\{X_i\}_{i=1}^N$  without any corresponding target variable. It is usually used for dimensionality reduction, information extraction, data visualization, density estimation, process monitoring and outlier detection (Ge et al., 2017). There are many unsupervised learning techniques such as K-means clustering, Principal component analysis (PCA), self-organizing map (SOM). In this research, PCA (Abdi & Williams, 2010; Dastile et al. 2020) was used to consider the accumulative effect in MMS, as discussed in the quality model section.

Supervised learning is applied when the dataset contains training samples  $\{x_i\}_{i=1}^N$  with the corresponding target variables  $\{y_i\}_{i=1}^N$ . The supervised learning can be formulated as

classification or regression problems based on the type of output variable. For a continuous output variable, regression algorithms such as linear, polynomial and support vector regression can be used; otherwise, classification algorithms can be used. Supervised learning is usually used for process monitoring, fault diagnosis, remaining useful life (RUL) estimation and quality prediction (Ge et al., 2017). In this research, random forest (Breiman, 2001), neural networks (Kotsiantis, 2007), K-nearest neighbors (Cunningham & Delany, 2007), support vector machine (Tsai et al., 2009), logistic regression, and Naïve Bayes algorithms were used to predict products' quality state (Kotsiantis, 2007).

### Quality monitoring using machine learning Techniques

Quality prediction and monitoring models that use different ML techniques have been applied in many manufacturing sectors such as semiconductor manufacturing (Al-Kharaz et al., 2019; Arif et al., 2013a, b; Kang et al., 2009; Kao et al., 2017; Melhem et al., 2015; Moldovan et al., 2017; Munirathinam & Ramadoss, 2016; Salem et al., 2018), steel manufacturing (Cuartas et al., 2020; Li et al., 2018; Lieber et al., 2013), 3D printing (Amini & Chang, 2018; Baturynska & Martinsen, 2021; Li et al., 2020), extrusion processes (García et al., 2019), battery manufacturing (Thiede et al., 2019), and automotive industry (Peres et al., 2019). The quality control and improvement using ML can be divided into four main tasks: description of product and process quality by identifying and ranking the most significant variables and factors related to the quality, prediction of the quality, classification of quality and optimization of the manufacturing parameters (Köksal et al., 2011).

The quality prediction models can be formulated as regression (second quality task) and classification (third quality task) problems based on output variable type. For continuous variables, regression models can estimate the quality characteristics of the intermediate and final products (Baturynska & Martinsen, 2021; García et al., 2019; Kang et al., 2009; Li et al., 2018; Thiede et al., 2019). Li et al., (2018) had successfully applied ML methods such as LR, SVM, K-NN and ensemble models by combining the used models with averaging and stacking methods to predict steel quality properties, e.g., tensile strengths, elongation. Thiede et al. (2019) have successfully applied multivariable regression to predict the characteristics of batteries, such as capacity. Kang et al. (2009) used virtual metrology (VM) based on regression models to estimate metrological values for each wafer based on manufacturing data and previous metrological results in an etching process of semiconductor manufacturing.

For discrete variables, classification models can predict the state of the product quality (conforming or

nonconforming) (Al-Kharaz et al., 2019; Amini & Chang, 2018; Arif et al., 2013a, b; Kao et al., 2017; Lieber et al., 2013; Malaca et al., 2019; Melhem et al., 2015; Peres et al., 2019). Lieber et al. (2013) have integrated supervised and unsupervised learning to predict steel quality and found that applying data mining had a significant effect on energy and time saving, especially in multistage manufacturing systems. Melhem et al. (2015) introduced a wafer quality prediction approach based on the health indicators' historical dataset. They proposed using PCA first to extract significant health indicators, then applying these indicators to the training process. The proposed approach has an acceptable performance. However, the health indicator extraction process could not be generalized to other applications.

Papers that applied ML algorithms for quality prediction and monitoring in MMS were rarely found. Arif et al. (2013a, b) introduced a cascaded quality monitoring framework to monitor the product quality in MMS. They used PCA to transform intercorrelated variables into a new set of variables that carry MMS characteristics then applied the iterative Dichotomiser (ID3) algorithm to predict the wafer's quality. The proposed approach has an acceptable false alarm rate, but the sensitivity was low, which limits the application in the quality field. Kao et al. (2017) had combined PCA with different classification techniques such as naïve Bayes, SVM and decision trees to predict the final product quality in the MMS; had used associated rule mining to find the root causes of defective products. Amini and Chang (2018) introduced an approach to predict the likelihood of printing non-defective products during 3D metal printing processes, especially at critical layers. They used K-means algorithms and random forest to determine the most critical layers. The proposed approach ignored the accumulative and interactive effects of printed layers.

### Challenges affect building an effective quality monitoring for multistage manufacturing system

In this section, we discuss four of the major challenges that affect building an effective quality monitoring model in MMS. The first is the curse of dimension, which can be handled by feature selection and extraction techniques. The second is the insufficient instances or samples of minority (defective) class samples which can be handled by using resampling techniques. The third is the appearance of missing data, which can be handled by using data imputation techniques. The fourth is the integration between the predictive model and quality inspection. Other challenges such as the accumulative and interactive effect of different workstations in the manufacturing chain and the changes in data characteristics and appearance of new patterns will be discussed in the methodology section.

## Feature selection and extraction

Feature selection is the process of determining a subset of variables or features that are most relevant to the quality prediction model. It simplifies the learning process, improves model performance, reduces computational time, and avoids overfitting. Many methods are used for feature selection, such as expert-based method, filter approach, wrapper approach, and embedded method (Colaco et al., 2019).

In filter methods, features are selected by calculating a score for each feature using a specific statistical criterion. Filter methods do not depend on a specific machine learning algorithm. Filter methods include Fisher discriminant metric, ANOVA, information gain and correlation coefficient. Filter methods have less computational burden compared to other feature selection methods, but they do not consider the interactions between features and have lower performance compared to other methods (Venkatesh & Anuradha, 2019).

Wrapper methods select a subset of features that maximize the performance of a specific learning algorithm. Wrapper methods include forward selection, backward elimination, genetic optimization and particle swarm optimization (Colaco et al., 2019). Wrapper Methods have higher performance than filter methods and consider the correlation between features, but they have higher computational burden compared to other methods and risk of overfitting (Venkatesh & Anuradha, 2019).

Embedded methods combine the characteristics of both filter and wrapper approaches. The selection of the feature is part of the model construction. Embedded methods include ridge regression and LASSO regression (Tadist et al., 2019). Embedded methods have less computational burden compared to wrapper methods and better performance than other methods, but they cannot work well for high dimensional data and have lower generalizability.

## Re-sampling methods

The insufficient samples of defective (nonconforming) products constitute a major challenge while building quality classification models. Since failure instances do not happen frequently, the number of positive class samples (nonconforming products) is smaller than the negative class samples (non-defective products). When a large majority or a small minority exists, the learning model will be forced to prefer the majority voice of the data to attain higher accuracy — This problem is called the imbalance problem —. Re-sampling methods can be used to rebalance the dataset. Re-sampling techniques are classified into oversampling and undersampling.

Undersampling techniques randomly remove samples from the majority class (conforming) samples, while oversampling increases the number of minority class

(nonconforming) samples so that the two classes become containing the same number of samples (Kotsiantis et al., 2006). The oversampling methods usually perform better than undersampling as removing samples increases the risk of losing useful information. However, the oversampling methods have higher learning time and risk of overfitting. Oversampling methods include techniques such as the synthetic minority over-sampling technique (SMOTE) (Chawla, 2009), duplicating the minor class samples and adaptive synthetic (ADASYN) (He et al., 2008).

## Missing value imputation methods

Missing data is a common problem in quality monitoring due to many causes, such as the low measurement and inspection rate due to limited capacity and time-consuming of the measuring equipment, and the errors in machines or measurement equipment (Yugma et al., 2015). Missing data can cause bias in estimating model parameters and loss of information (Lee et al., 2019). There are three types of mechanisms in which missing values can occur. Missing completely at random (MCAR) where the probability of missing values of a variable is unrelated to the values of the variable itself and other measured variables — do not depend on unobserved or observed data —. Missing at random (MAR), Where the missing values are unrelated to unobserved data but related to other measured or observed variables. MAR is actually not random but it describes a systematic missing values where missing values is correlated with other associated variables in the analysis. Missing not at random (MNAR) when the probability of missing values in a variable is depending on the variable itself (Baraldi & Enders, 2010).

Imputing missing data include deletion method, Single imputation methods such as mean, stochastic regression imputation and regression imputation, and Multiple imputation techniques such as Markov chain Monte Carlo (MCMC) method and expectation maximization (EM) algorithm (Nakagawa, 2015).

## Implementation of the predictive model in quality monitoring

Most of the quality prediction models in the literature (Sect. 2.2) show promising results in terms of model performance; however, the integration between the prediction models and the quality inspection and planning or the implementation in quality activities was rarely found. Schmitt et al. (2020) introduced a technical implementation of predictive models in quality inspection. They found that using predictive models in the quality inspection had significantly reduced inspection volumes. Kang and Kang (2017) had introduced intelligent virtual metrology (VM) system based

on adaptive updates using an ensemble ANN to reduce the metrology cost and achieve superior performance.

## Quality monitoring framework

In this section, we introduce a smart real-time quality monitoring system that integrates supervised and unsupervised ML techniques to predict the quality deviations for complex multistage manufacturing processes and introduce a hybrid quality inspection strategy based on predictive models and physical inspections to reduce inspection costs and time and achieve zero defect manufacturing. The proposed framework consists of two primary phases: Phase I is the data preprocessing and model building phase; phase II is monitoring the products, implementing the hybrid quality inspection strategy and updating the monitoring model.

### Phase I: Data gathering and model building

Phase I aims to build a quality monitoring and inspection system capable of overcoming the first five challenges while building the quality monitoring model for MMS. This phase consists of several procedures related to data gathering and preprocessing, building the quality prediction model with considering relationships between workstations in MMS and the unbalanced nature of the manufacturing process, and determining the most important variables and significant manufacturing stages in the manufacturing chain.

#### Data gathering

Smart sensors and other data collection techniques gather data and measurements from different sources along the manufacturing chain. The gathered data includes maintenance records, state variables, process parameters, intermediate product characteristics, final product quality and environmental sensing variables (Turetsky et al., 2020). The quality-related data can be determined based on expert knowledge and previous manufacturing experiment.

#### Data preprocessing

The manufacturing environment contains many anomaly events, outliers, noises, missing and inconsistency values. Raw data needs to be preprocessed carefully to remove misleading, redundant and inconsistent information. Data Preprocessing improves the quality of the gathered raw data and enhances model performance (Xu et al., 2015). Data preprocessing involves data cleaning and transformation.

The first step of data cleaning is to remove outliers and gross errors as they mislead and disturb the training process and reduce model performance.  $3\sigma$  rule or IQR (interquartile

range) can be used to remove outliers and gross errors. Secondly, variables and instances that contain missing values above a tuned thresholds P1 and P2 needed to be removed. Third, the remainder missing values are substituted by attribute mean as attribute mean has low computational effort and can express initial estimation of the model performance and K-NN imputation for the best model as K-NN can deal with the three mechanisms of missing data and has an appropriate performance. Fourth, columns with constant and duplicated values needed to be removed as they do not contain any distinctive information (Xu et al., 2015). After cleaning the dataset, data transformation is required to adjust the scale difference between the manufacturing variables. Data are normalized between zero and one using Eq. (1). Finally, the preprocessed dataset is split into 80% and 20% for training and testing.

$$Y = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

where  $X_{\min}$  and  $X_{\max}$  are minimum and maximum values for each variable.

#### Feature selection

In this research, The AUC-based permutation feature importance measure is used, as it is more robust against imbalanced datasets (Janitza et al., 2013). The AUC score of the out-of-bag (OOB) instances is calculated before and after randomly shuffling the values of a predictor variable. The main idea of this method is that if there is an association between a variable and the response, breaking this association by randomly shuffling the values of this variable will influence the performance of the model. The variable importance scores can be calculated by averaging the difference in AUC score over the number of trees in the forest, as shown in Eq. 2 (Janitza et al., 2013).

$$VI_j^{AUC} = \frac{1}{n_{tree}} \sum_{i=1}^{n_{tree}} (AUC_{ij} - AUC_{\hat{j}}) \quad (2)$$

where  $VI_j^{AUC}$  is the variable importance,  $n_{tree}$  is the number of trees in the forest.  $AUC_{ij}$  and  $AUC_{\hat{j}}$  are the area under the ROC calculated from the OBB instances in the  $i^{th}$  tree before and after randomly permuting variable  $j$ .

#### Tackling the imbalance problem

Synthetic Minority Over-sampling Technique (SMOTE) is used in this research to rebalance the training dataset as it has lower computational effort, especially for high dimensional data, and has appropriate performance (Lee et al., 2019). SMOTE increases the number of defective class samples and

makes them equal to the number of non-defective class samples by generating new samples using the K-NN concepts (Chawla, 2009). SMOTE can be summarized as for a sample  $x$  in the minority (defective) class, compute the Euclidean distance between this sample and all samples in the minority class, then list it  $K$  nearest neighbors based on the calculated distance. Based on the required percentage of oversampling, random samples  $z_i$  ( $i = 1$  to  $k$ ) from its  $K$  nearest neighbors are selected and for each  $z_i$  a new sample  $S$  is synthesized using the Eq. (3). the previous steps are repeated until the balancing is achieved.

$$S = x + a(z_i - x) \tag{3}$$

where  $a$  is a random variable between  $[0,1]$ .

**Cascade qquality model**

In MMS, the quality of the produced product from a manufacturing stage is influenced by the manufacturing variables of this stage and the quality characteristics of the produced product from the former stage; thus, each manufacturing stage affects the final quality of the products (Arif et al., 2013a, b). Partial and total quality principles needed to be introduced to represent the accumulation of variations and the relations between different manufacturing stages along the manufacturing chain (Zhai et al., 2004). Cascade Quality Prediction Method (CQPM) (Ge et al., 2017; Kim et al., 2012) is used to describe and analyze the relationships in MMS. There are three kinds of relationships in MMS:  $R_1$  is the relationship between manufacturing variables in each manufacturing stage,  $R_2$  is the relationship between different manufacturing stages in the MMS, and  $R_3$  is the relationship between manufacturing variables and the final product

quality. Figure 1 shows the relationships in MMS and partial and total quality concepts.

Based on the CQPM approach (Arif et al., 2013a, b), it was assumed that the manufacturing chain consists of  $M$  manufacturing stages. Each stage has several manufacturing operation variables  $X_s$  and each intermediate product has many quality characteristics  $T_s$ . PCA was used to transform the intercorrelated and latent manufacturing variables into a new set of dimensions (Principal components) that carry the characteristics of MMS. Equation 4 shows how the interaction between manufacturing variables ( $X_{j,k}$ ) and the quality characteristics  $T_{(j-1),k}$  from the former workstation can influence the quality of intermediate products, which expresses  $R_1$  and  $R_2$ . Equation 5 shows how the quality characteristics from the final manufacturing stage ( $T_{M,s}$ ) influence the quality state of the final product quality  $Q$ , which expresses  $R_3$ .

$$T_{j,s} = f(T_{(j-1),s}, X_{j,k}) = \sum_{k=1}^{P_{(j-1)}} a_{s,k} T_{(j-1),s} + \sum_{k=1}^{P_j} a_{s,k} X_{j,k} \tag{4}$$

$$Q = f(T_{M,s}) \tag{5}$$

where:  $X_{j,k}$  is the  $k$ th manufacturing variable in  $j$ th manufacturing stage ( $j = 1, 2, \dots, M$  and  $K = 1, 2, \dots, p$ );  $T_{j,s}$  is the  $s$ th quality characteristic of the intermediate product produced from the  $j$ th manufacturing stage ( $s = 1, 2, \dots, r$ );  $a_{s,k}$  is the amount of contribution of  $X_{j,k}$  to  $T_{j,s}$  (the eigenvector of the covariance matrix in PCA) and  $Q$  is the quality level of the final product.

**Building the quality prediction model**

The selection of an adequate supervised ML algorithm at the beginning cannot be achieved in a general way. Different

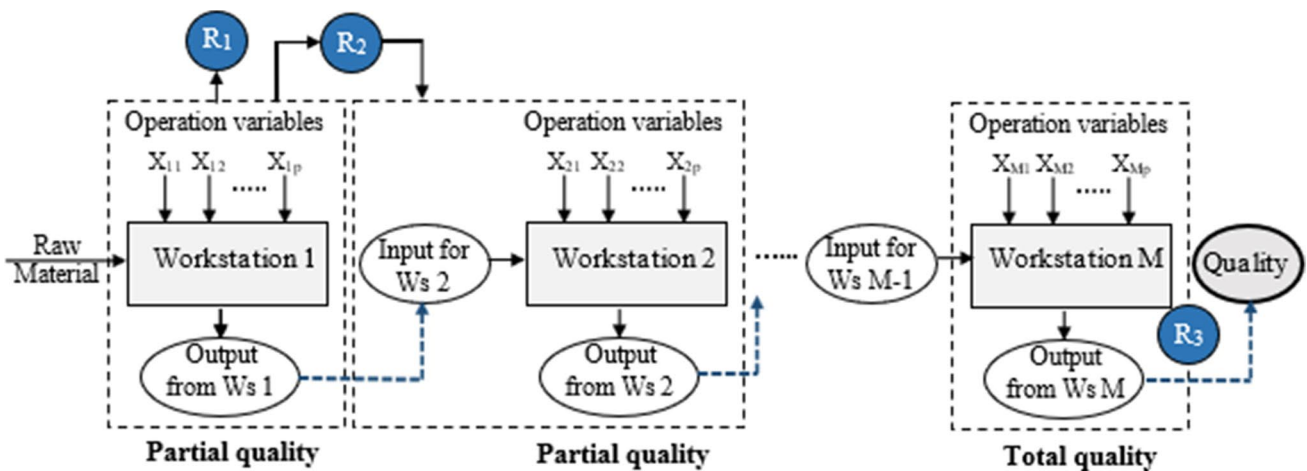


Fig. 1 Relationships in MMS and partial and total quality principals

classification algorithms should be trained and evaluated for each application to choose the best performing one (Lee & Shin, 2020). The models’ preselection should be based on selected criteria, e.g., speed, complexity and interpretability, or the experience of data scientists and results from similar or previous projects. Several classification algorithms such as SVM, K-NN, naïve Bayes ANN, logistic regression and random forest should be applied and evaluated in order to select the best algorithm for the quality prediction. K Fold cross-validation and grid search were used to tune the hyper-parameters of the predictive models and to avoid overfitting (Raschka, 2018).

Different performance metrics such as accuracy, specificity, sensitivity, precision, negative predictive value and AUC can be used to evaluate and compare the classification models. These metrics can be calculated from the confusion matrix, as shown in Fig. 2 and calculated by using equations from 6 to 10 (Tharwat, 2018).

Two types of errors occur during the prediction and inspection process. Type I error occurs when the prediction model or the inspection process wrongly rejects a conforming (non-defective) product. In contrast, Type II error occurs when the prediction model or the inspection process wrongly accepts a nonconforming (defective) product (Sarkar & Saren, 2016). False positive (FP) and False negative (FN) that are shown in the confusion matrix (Fig. 2) represent Type I and Type II errors, respectively. Type II error is more serious for most companies as its effects are touched directly by customers, leading to losing customers’ trust.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

$$specificity = \frac{TN}{TN + FP} \tag{7}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

		Actual class	
		Nonconforming product	Conforming Product
Predicted class	Nonconforming product	True Positive (TP)	False Positive (FP)
	Conforming Product	False Negative (FN)	True Negative (TN)

Fig. 2 Confusion matrix

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Negative\ prediction\ value = \frac{TN}{TN + FN} \tag{10}$$

### Stage selection

In MMS, not all the manufacturing stages have an equal influence on the final product quality. So, it is suggested to determine the most influencing manufacturing stages on the final product quality then build quality checkpoints at these critical stages to predict the likelihood of producing non-defective products instead of checking and predicting the quality only at the final manufacturing stages. The conventional methods for identifying the critical stages are based on the knowledge and experience of engineers. Currently, the manufacturing chain can contain more than 500 steps (e.g., semiconductor manufacturing) (Lee et al., 2019). In this research, the critical stages were identified by summing the variables’ importance values computed from the AUC-based permutation feature importance for each manufacturing stage. After that, a significant level was set to identify the significant stages. Identifying the critical steps can help to enhance the production yield and reduce waste of resources and time.

### Phase II: Process monitoring and implementing the hybrid quality inspection strategy

This phase aims to monitor the product quality along the manufacturing chain, make informed decisions about the product’s quality in real-time and cope with the last two challenges that affect building an effective quality monitoring model in MMS. This phase consists of three steps: monitoring the intermediate product quality at the critical manufacturing stages and final product quality at the end of the production line in order to take appropriate actions to save resources; implementing the quality prediction in quality inspection and planning through hybrid quality inspection strategy; updating the monitoring model to cover any data characteristics change.

### Online quality monitoring

CQPM is implemented to the manufacturing variables in each manufacturing stage and quality characteristics from the former stage. When the first important stage is attained, the first prediction model will give an initial estimation of the final product quality. If the likelihood of prediction is -1 (conforming product), the manufacturing process continues until the following significant stage is attained. Otherwise,

the process engineer or an automated adjustment system can take preemptive action to save time and resources or adjust the manufacturing parameters in the following stages so that the intermediate product may have a chance of being good at the final stage. When the final stage is reached, the final product quality can be estimated.

### Hybrid quality inspection strategy

When a product is produced by certain manufacturing equipment, a small number of products are inspected according to the quality inspection protocol using destructive or non-destructive inspection (Filz et al., 2020; Foumani & Tavakkoli-Moghaddam, 2019; Li et al., 2018; Mineo et al., 2016). This inspection approach assumes that the manufacturing conditions of each produced product are the same and the manufacturing errors between samples are identically distributed and statistically independent (Lee, 2020). Due to the high number of variables that needed to be monitored and controlled in MMS, abnormalities usually appear. So, the sampling approach may not fully reflect the actual quality, especially in complex manufacturing processes. As the final product's quality state based on the prediction model only requires a too high confident model that exceeds or at least reaches the conventional inspection level, the hybrid inspection strategy seems promising for the MMS. Thereby, the inspection strategy will be given by the combination of conventional inspection and quality prediction models. It is suggested to apply physical inspections only on the products predicted to be defective by the predictive model. The hybrid inspection strategy requires tuning the prediction model with respect to zero false negatives (FN) to reduce Type II error. The reason for tuning the model with respect to zero FN is that the risk and cost of Type II error are usually higher than Type I error (Sarkar & Saren, 2016). This promising strategy will save inspection time, cost and ensure the consistency of the output quality.

### Model updasting

The monitoring model should be updated periodically to handle any new patterns and any changes in the data characteristics (Hirai & Kano, 2015). The changes occur due to some expected and unexpected events such as maintenance, change in loads and internal and external disturbances (Kang & Kang, 2017). These changes make monitoring models unreliable. So the quality monitoring model should be updated to reduce the chance of performance degradation and misclassification. The updating of the model can be made after some known events such as maintenance; however, the change can sometimes happen due to some unknown events. So, it is suggested to use a critical threshold to give a warning when the performance decreases and

the misclassification increases. Hence, the quality monitoring model can be updated by retraining the model using the newly collected data from the physically inspected products plus the initial dataset. The quality monitoring system is summarized in Fig. 3.

### Cases study

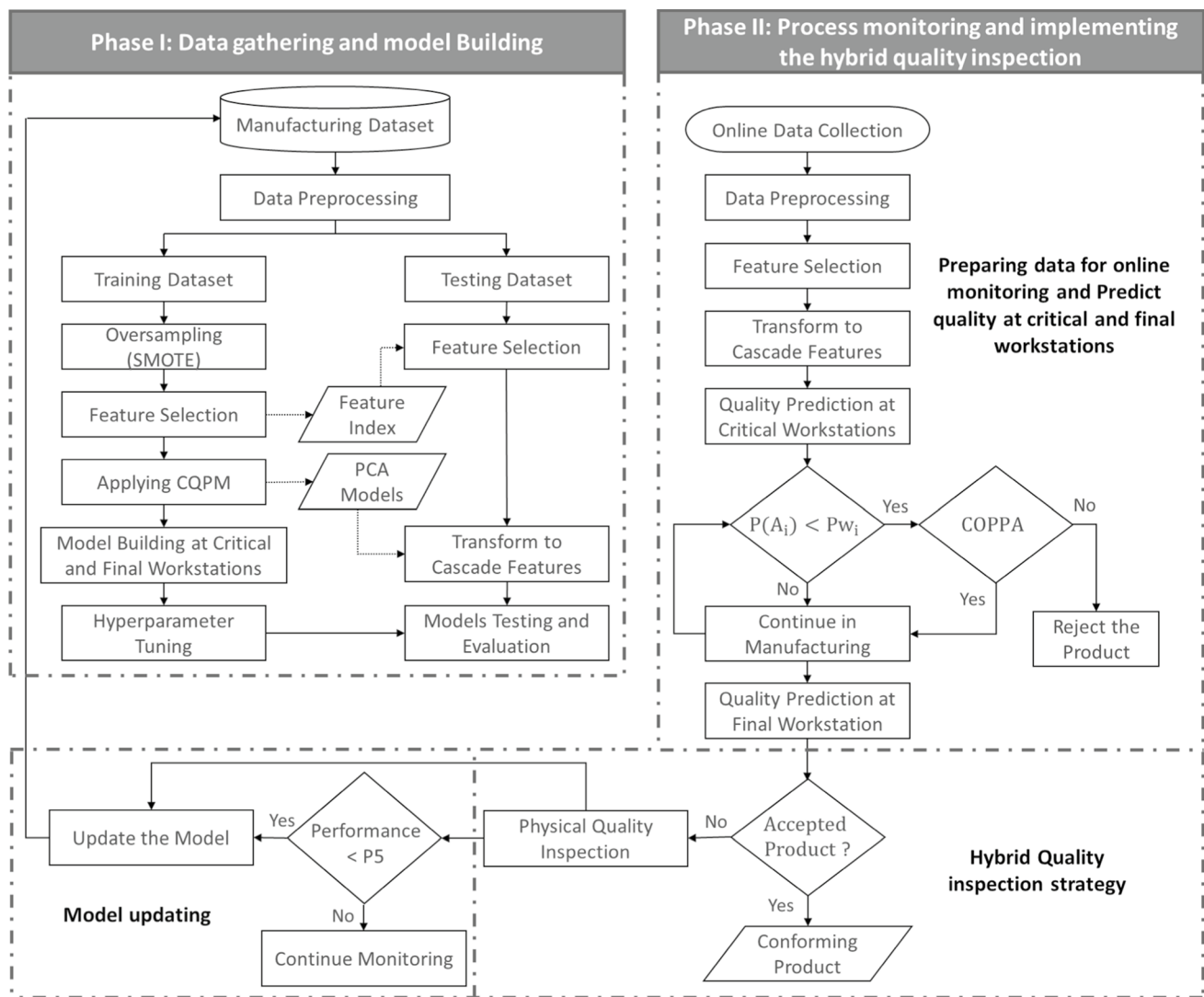
A complex semiconductor manufacturing dataset (called SECOM) was used to verify the proposed framework (McCann & Johnston, 2008). This dataset contains 1567 samples of semiconductor wafers and 590 manufacturing variables gathered by different sensors along the semiconductor wafer production line. Each sample has one quality index, which is labeled with -1 for non-defective wafers and 1 for defective wafers; each sample is associated with a particular timestamp. The dataset is highly imbalanced as there are 104 defective wafers and 1463 non-defective wafers (1:14 approximate ratio between the two classes). The dataset contains 41,951 (4.6%) missing values distributed with different percentages in manufacturing variables and 116 irrelevant and constant values variables.

The dataset was divided into five groups; each group represents a workstation to emulate the MMS scenario. Two quality monitoring systems were built based on the proposed framework. The first system is based on the suggested framework and the second system is based on single-point approaches where the whole dataset was treated as a single point. The two systems were compared using different metrics to evaluate the capability of the suggested system for quality monitoring in MMS. The most influencing stages were identified and used as quality checkpoints. Also, the hybrid quality inspection strategy was successfully implemented. All the computations were conducted by using R Studio (RStudio, 2020).

### Building the quality monitoring model

To handle this complex dataset, first, the data was cleaned by removing variables and samples containing a percentage of missing values higher than thresholds  $P1 = 40\%$  and  $P2 = 25\%$ . Some papers suggested using a threshold of 37% or 55% for the same dataset (Munirathinam & Ramadoss, 2016; Salem et al., 2018), but it was found that our thresholds perform better for our proposed framework. No samples were removed from the dataset as the performance decreased when we removed instances at different thresholds. The Irrelevant and low variance features were removed from the dataset as they do not hold any distinctive information. Outliers and gross errors were removed using the  $3\sigma$  rule as they reduce the model performance and increase the training time. The former steps helped to decrease the dimension





**Fig. 3** summary of the Quality monitoring framework.  $P(A_i)$ : Probability of producing a conforming product.  $P_{w_i}$ : a threshold for probability of producing a conforming product at critical manufacturing stages. COPPA, the chance of process parameter adjustment in the

following workstation to produce a good product. CQPM: Cascade quality prediction method.  $P_5$ : a threshold for updating the training dataset

of data to 431 features. The rest of the missing values were substituted by the attribute mean for each variable. The pre-processed data were normalized between zero and one using Eq. 1. The dataset was randomly divided into 80% and 20% for training and testing. The training dataset was rebalanced by using a re-sampling technique called SMOTE, so that the proportion between the defective and non-defective samples in the training dataset is approximately equal to 1. The AUC-based permutation feature importance measure (Janitza et al., 2013) was used to order the manufacturing variables based on their importance to the quality of the semiconductor wafers and remove variables that are less relevant to the quality in order to simplify the learning process, reduce computational time, avoid overfitting and improve

model performance. Figure 4 shows the performance at a different number of features. According to Fig. 4, the performance increases with increasing the number of features, and then it decreases with a small fluctuation in performance of some metrics after reaching the optimum number of features. It was found that using the highest 65 important features gives the best performance in terms of sensitivity and AUC and an acceptable specificity value (Fig. 4). The previous preprocessing steps helped to reduce the dimension of data from 590 to 65 variables.

In order to emulate the MMS scenario, the dataset was divided into five groups; each group represents a manufacturing stage as recommended by (Kim et al., 2012). Figure 5 represents the 65 top selected features

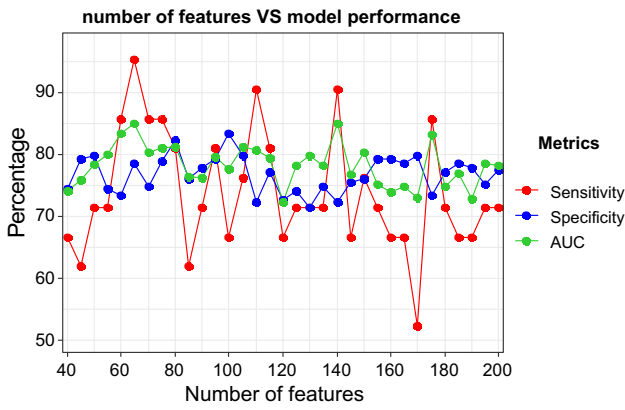


Fig. 4 Model performance at a different number of features

distributed over the five workstations. Each workstation’s accumulative and interactive effect was considered by applying the CQPM based on PCA to each workstation’s manufacturing variables and the previous workstation’s quality characteristics. Principal components that represent 95% of the variance for manufacturing variables of each manufacturing stage and quality characteristics from the previous stage were chosen. Table 1 shows the number of principal components that represent the accumulative effect for each workstation.

Six ML algorithms were applied, namely K-NN, Naïve Bayes, random forest, SVM, ANN and logistic regression, to estimate the quality of the wafers. tenfold cross-validation and grid search were used to avert overfitting and for hyperparameter tuning for each ML model.

Table 1 Principal components that represent 95% of the variance and the accumulative effect at each workstation

	Workstation 1	Workstation 2	Workstation 3	Workstation 4	Workstation 5
Principal Components (Mean Imputation)	9	17	23	24	26
Principal Components (K-NN Imputation)	10	18	25	29	31

### Evaluating and enhancing the capability of the proposed system

The suggested quality monitoring system (first system) was compared with the single-point approach to ensure the capability of the suggested system for quality monitoring in MMS. The single-point approach was built on the same proposed framework, but instead of using CPQM, the whole dataset was treated as a single point. The number of principal components for the single point approach was 35 and it represents 95% of the variance in data. Figure 6 shows the comparison between the two quality monitoring approaches. The comparison between the two approaches was made using sensitivity, specificity and accuracy metrics. Sensitivity, specificity represents the ability to detect the

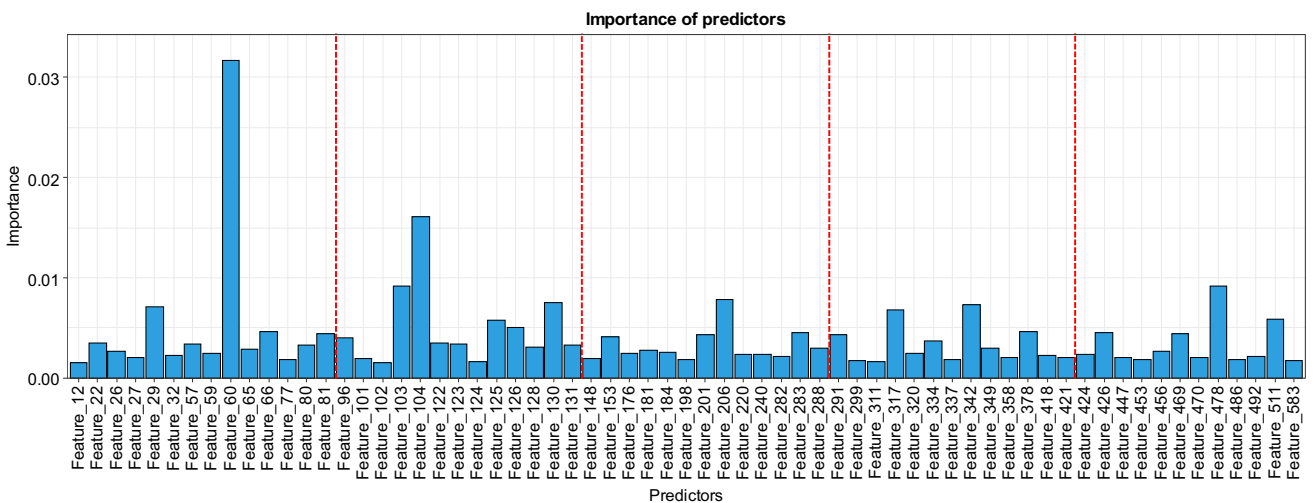
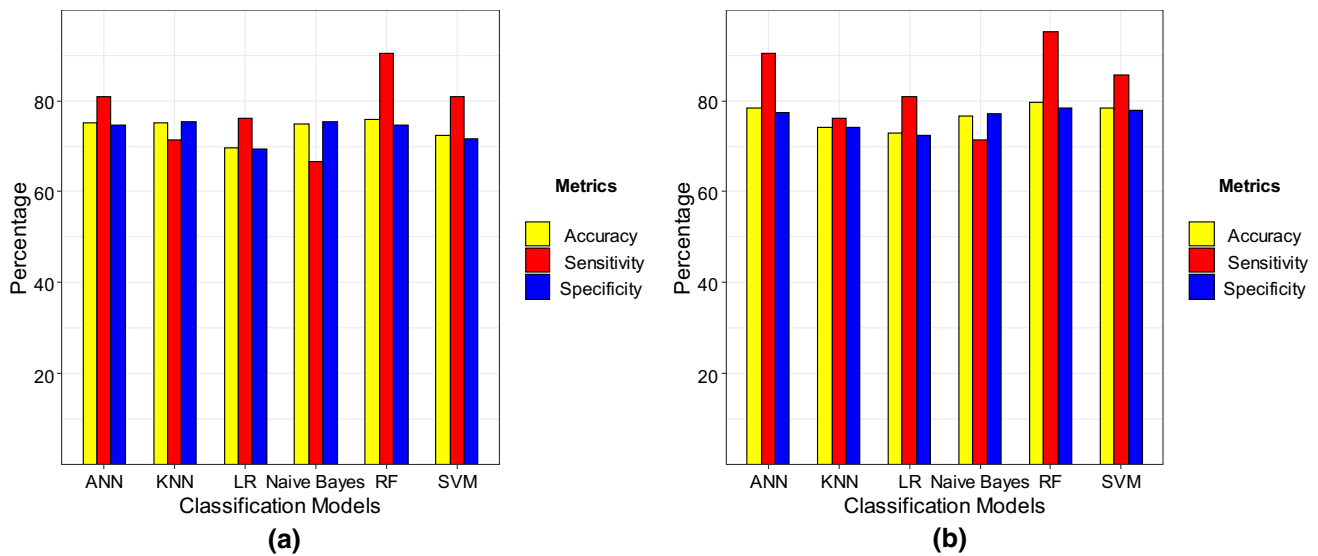


Fig. 5 The distribution of the 65 top importance features over the five workstations



**Fig. 6** Performance of the quality monitoring models using a) single point quality monitoring approach and b) the proposed framework (first system)

defective and non-defective wafers; also, specificity can refer to the false alarm rate, which are essential in evaluating the quality monitoring processes.

It was found that using the proposed system for MMS performed better than using the single point approach in terms of accuracy, specificity and sensitivity. The average capability of detecting the nonconforming product for the six models, which can be described by sensitivity, has improved by 5.6%. The average false alarm rate for the six models has reduced by 2.7% and the average accuracy has improved by 2.9%. The results show the ability of the proposed model to represent the accumulative and interactive effect of each manufacturing stage and the ability to enhance the quality monitoring in MMS.

With the focus on the quality monitoring for the first system; accuracy, sensitivity, specificity, precision, negative prediction value and AUC were used to evaluate and compare the performance of different quality predictive models. Table 2 represents the performance of classification models on the test set.

The random forest algorithm performed the best between all classification models (Table 2). With sensitivity and

specificity of 0.952 and 0.785, respectively, most of the defective wafers were detected correctly with a 21.5% false alarm rate. ANN had the highest AUC score, but in overall, the random forest was the best classification algorithm as it has the highest sensitivity and specificity, precision and negative prediction value. The strength of the RF model comes from merging the idea of bagging with bootstrap samples and random subspace. This merging reduces and smooths the variance by building multiple and uncorrelated trees. Those trees spread potential errors, cancel them out through the majority voice and make the model predict results that are less away from the actual values.

In order to enhance the random forest model performance in terms of specificity and precision, K-NN was used for missing data imputation instead of mean imputation (García-Laencina et al., 2010). The previous steps for the model building were repeated and the random forest algorithm was re-evaluated. The optimum number of features when using K-NN for missing value imputation was 70 features and the distribution of features along the manufacturing chain is shown in Fig. 7. Table 1 shows the number of principal components that represent the accumulative effect for

**Table 2** Performance for different quality monitoring models

Evaluation Metrics	Naive Bayes	K-NN	RF	SVM	LR	ANN
Accuracy	0.768	0.742	0.796	0.783	0.729	0.783
Sensitivity	0.714	0.762	0.952	0.857	0.810	0.905
Specificity	0.771	0.741	0.785	0.778	0.724	0.775
Neg Pred Value	0.974	0.977	0.996	0.987	0.981	0.991
Precision	0.183	0.174	0.241	0.217	0.173	0.224
AUC	0.747	0.794	0.849	0.796	0.802	0.873

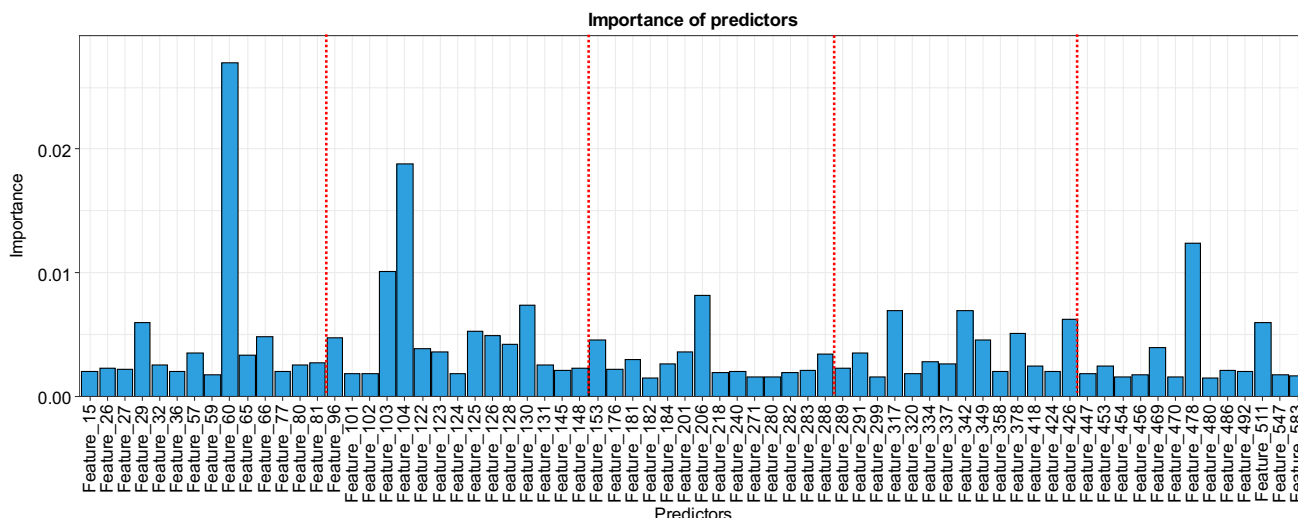


Fig. 7 the distribution of the 70 top importance features over the five workstations

Table 3 Random forest performance after using K-NN for missing data imputation

Model	Accuracy	Sensitivity	Specificity	Precision	Neg Pred Val	AUC
Random Forest	0.815	0.952	0.805	0.26	0.996	0.86

Table 4 The importance of each workstation

	Workstation 1	Workstation 2	Workstation 3	Workstation 4	Workstation 5
Impact of each workstation	0.239	0.278	0.147	0.187	0.15

Table 5 Performance at significant manufacturing stages

Significant stages	Accuracy	Sensitivity	Specificity	Neg Pred Value	Precision
Stage 1	0.694	0.619	0.7	0.962	0.129
Stage 2	0.723	0.762	0.72	0.977	0.163

each workstation. Using K-NN for missing data imputation helped to improve specificity and precision by 2% and 1.9%; other metrics except sensitivity had also improved. Table 3 shows the performance of random forest when using K-NN for missing value imputation.

### Identifying the most influencing manufacturing stages

The critical stages can be identified by summing the variables’ importance values computed from the AUC-based permutation feature importance for each workstation, as it is shown in Fig. 7. A significant level of 0.2 was chosen to select the most critical workstation. Table 4 shows the importance of each workstation. Based on the significant level, the first and second workstations were identified as critical workstations. Hence, quality checkpoints were built at those important stages to provide an initial estimation

about the intermediate product quality and at the final stage to estimate the final product quality state.

As the random forest performance was better than other classification algorithms, it was used to predict the likelihood of producing non-defective products at the critical manufacturing stages and provide an initial estimation of the final product quality. The performance of estimating the final product quality of the intermediate products at the critical stages (first and second manufacturing stages) is shown in Table 5.

If the likelihood of prediction at critical stages (stage 1,2) is -1 (conforming product), the manufacturing process continues until the following significant stage is attained. Otherwise, the process engineer or an automated adjustment system can take preemptive actions to save time and resources or adjust the manufacturing parameters in the following stages so that the intermediate product may have a chance of being good at the final stage. The performance of all metrics at the second quality checkpoint increased due to the addition of more

influencing variables from the second manufacturing stage to the quality characteristics from the first stage.

### The hybrid quality inspection strategy

When the product reaches the final manufacturing workstation, the hybrid inspection approach can be implemented. As the negative predictive value is 0.996 as shown in Table 3, most of the conforming wafers (99.6% of the wafers) were correctly predicted by the quality prediction model to be non-defective. Instead of inspecting a few samples of wafers from each lot according to the inspection protocol or inspecting all the produced wafers, which are time and cost consumption, or using only prediction models. Only those wafers that were predicted to be defective will be inspected. We will suffer from false alarms as the precision = 0.26, which means that 26% of the predicted products are truly defective while the others are non-defective. Nevertheless, the hybrid strategy still has promising results as physical inspection was reduced by 75.2%. The inspection time and cost will significantly be reduced due to the reduction in inspection volume; also, the confidence in the quality of the produced product will be increased.

There are some complex inspection scenarios that can happen in the actual manufacturing environment. One scenario includes the limited inspection allowance where products cannot be reworked for an infinite number of times until it succeeds, inspection process with decreasing failure rate and limited inspection time. We refer the readers to Foumani et al. (2020) and Rezaei-Malek et al. (2019) for more about those complex scenarios.

The training dataset should be retrained and updated periodically to handle any new patterns and any change in data characteristics to reduce the chance of performance degradation and misclassification. So, it is suggested to use a critical threshold based on the precision or false alarm rate to give a warning when the performance decreases and the misclassification increases. Hence, the quality monitoring model can be updated by retraining the model using the newly collected data from the physically inspected products plus the initial dataset.

### Ablation study

We implemented an ablation study on the proposed framework to identify the contribution of different steps and components to the overall performance (Lipton & Steinhardt, 2019). The ablation study was performed using the random forest algorithm because its performance was better than the other classification algorithms. Three models were conducted based on the proposed framework; the first model was constructed by using all pre-processed features without considering the problem of high dimensional data (except the feature selection step). The second model was built without balancing the training dataset or without considering the imbalanced nature of the manufacturing processes (except the resampling step). The third model was constructed without CQPM. In order to have a fair comparison, the same steps for model building and evaluation (in phase I) were used for the three models. Table 6 shows the results of the ablation study on the random forest algorithm.

It can be observed that removing one of the three previous steps causes degradation of the model performance. For the first model, sensitivity has reduced by 28.5% and specificity has not changed. For the second model, sensitivity and specificity have reduced by 28.5% and 15.5%, respectively, which was the highest performance degradation. The sensitivity has not changed for the third model, while the specificity has decreased by 6.8%. From the ablation study, we concluded that using the full model would be better for quality monitoring in MMS.

### Limitation of the study

In this study, it was assumed that the manufacturing chain consists of five workstations. However, in the real-world, the semiconductor manufacturing process consists of hundreds of steps to produce a wafer. This requires enormous data and high computational effort. In such a case, the quality monitoring model can be built by using the most important stages (Amini & Chang, 2020; Lee et al., 2019). Using the most important stages may cause losing some information that may be related to the final product quality. However, this can reduce computational time and effort. Also, in this research, we deal with quantitative data. However, the data can be gathered from different sources with different types (e.g. text data from manual inspections). So, various data

**Table 6** Ablation study on the proposed framework

Models	Accuracy	Sensitivity	Specificity	Neg Pred Value	Precision	AUC
Without feature selection	0.799	0.667	0.808	0.971	0.2	0.741
Without resampling	0.656	0.667	0.655	0.965	0.122	0.747
Without CQPM	0.748	0.905	0.737	0.991	0.198	0.814
Full Model	0.815	0.952	0.805	0.996	0.26	0.86

types should be considered, e.g., text mining methods for text data.

The effect of the suggested quality monitoring and inspection model on time, resources and cost-saving have not been evaluated as a secondary dataset was used. To overcome this limitation, a real case study can be used in the future. The optimum time to update the training data before the performance of the prediction degradation reaches the defined threshold needs more investigation and study.

## Conclusion and future work

This research introduces a smart real-time quality monitoring and inspection framework to predict and determine the quality deviations for complex and multistage manufacturing processes as early as possible. The proposed framework consists of two primary phases: Phase I is the data preprocessing and model building phase, which includes data gathering and preprocessing, feature selection, building the quality prediction model considering the accumulative effects of workstations and identifying the critical stage to the final product, while phase II includes monitoring products quality, implementing the hybrid quality strategy based on the prediction model and the physical quality inspection and updating the quality monitoring model to handle any new patterns to decrease the chance of misclassification.

The proposed framework was verified using a complex semiconductor manufacturing dataset. The quality monitoring model based on the proposed system and single-point approach was compared. The results show the ability of the proposed framework to enhance the quality monitoring performance in the multistage manufacturing processes. The average performance of the six classification models in terms of specificity and sensitivity has increased by 2.7% and 5.6%, respectively. The random forest algorithm performed the best between all classification model models with 95.2% sensitivity and 80.5% after using K-NN for missing value imputation. Stages 1 and 2 were identified as the most influencing stages on the final product quality and were used as quality checkpoints to provide an initial estimation of the final product quality. The hybrid quality inspection strategy based on predictive models and physical inspection is implemented when the product reaches the final workstation. It shows promising results in reducing inspection volume by 75.2%, also the inspection cost and time will probably be decreased.

There is a direct relation between inspection volume, cost, and time. The testing operation and physical metrology operations are usually expensive and time consuming, especially in some applications such as semiconductor manufacturing processes and steel production processes. Reducing the inspection volume will directly reduce both testing

operation and physical metrology operations cost and time, which will lead to reducing both inspection cost and time.

For future work, we plan to use more advanced data mining and preprocessing techniques to achieve nearly 100% negative predictive value, reduce the false alarm and the useless inspection of conforming products. We also plan to implement the proposed framework into a real case study to measure the system's impact on resources, cost, and time-saving. Finally, some challenges, such as model reliability, updating and the initial estimation of product quality, needed to be deeply investigated.

## References

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley. Computational Statistics*, 2(4), 433–459. <https://doi.org/10.1002/wics.101>
- Al-Kharaz, M., Ananou, B., Ouladsine, M., Combal, M., & Pinaton, J. (2019). Quality prediction in semiconductor manufacturing processes using multilayer perceptron feedforward artificial neural network. In: *8th international conference on systems and control (ICSC 2019)*, pp. 423–428. <https://doi.org/10.1109/ICSC47195.2019.8950664>
- Amini, M., & Chang, S. I. (2018). MLCPM: A process monitoring framework for 3D metal printing in industrial scale. *Computers and Industrial Engineering*, 124, 322–330. <https://doi.org/10.1016/j.cie.2018.07.041>
- Amini, M., & Chang, S. I. (2020). Intelligent data-driven monitoring of high dimensional multistage manufacturing processes. *International Journal of Mechatronics and Manufacturing Systems*, 13(4), 299–322. <https://doi.org/10.1504/IJMMS.2020.112352>
- Arif, F., Suryana, N., & Hussin, B. (2013a). Cascade quality prediction method using multiple PCA+ID3 for multi-stage manufacturing system. *IERI Procedia*, 4, 201–207. <https://doi.org/10.1016/j.ieri.2013.11.029>
- Arif, F., Suryana, N., & Hussin, B. (2013b). A data mining approach for developing quality prediction model in multi-stage manufacturing. *International Journal of Computer Applications*, 69(22), 35–40. <https://doi.org/10.5120/12106-8375>
- Aydemir, G., & Acar, B. (2020). Anomaly monitoring improves remaining useful life estimation of industrial machinery. *Journal of Manufacturing Systems*, 56, 463–469. <https://doi.org/10.1016/j.jmsy.2020.06.014>
- Bai, Y., Sun, Z., Zeng, B., Long, J., Li, L., de Oliveira, J. V., & Li, C. (2019). A comparison of dimension reduction techniques for support vector machine modeling of multi-parameter manufacturing quality prediction. *Journal of Intelligent Manufacturing*, 30, 2245–2256. <https://doi.org/10.1007/s10845-017-1388-1>
- Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology*, 48, 5–37. <https://doi.org/10.1016/j.jsp.2009.10.001>
- Baturynska, I., & Martinsen, K. (2021). Prediction of geometry deviations in additive manufactured parts: comparison of linear regression with machine learning algorithms. *Journal of Intelligent Manufacturing*, 32(1), 179–200. <https://doi.org/10.1007/s10845-020-01567-0>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chawla, N. V. (2009). Data mining for imbalanced datasets: An overview. In O. Maimon & L. Rokach (Eds.), *Data Mining and*

- Knowledge Discovery Handbook* (pp. 875–886). Boston, MA: Springer. [https://doi.org/10.1007/978-0-387-09823-4\\_45](https://doi.org/10.1007/978-0-387-09823-4_45)
- Cheng, Y., Chen, K., Sun, H., Zhang, Y., & Tao, F. (2018). Data and knowledge mining with big data towards smart production. *Journal of Industrial Information Integration*, 9, 1–13. <https://doi.org/10.1016/j.jii.2017.08.001>
- Cohen, Y., & Singer, G. (2021). A smart process controller framework for Industry 40 settings. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-021-01748-5>
- Colaco, S., Kumar, S., Tamang, A., & Biju, V. G. (2019). A review on feature selection algorithms. In: Shetty, N., Patnaik, L., Nagaraj, H., Hamsavath, P., Nalini, N. (Eds.), *Emerging research in computing, information, communication and applications. Advances in intelligent systems and computing*, vol 906. Springer, Singapore. [https://doi.org/10.1007/978-981-13-6001-5\\_11](https://doi.org/10.1007/978-981-13-6001-5_11)
- Cuartas, M., Ruiz, E., Ferreño, D., Setién, J., Arroyo, V., & Gutiérrez-Solana, F. (2020). Machine learning algorithms for the prediction of non-metallic inclusions in steel wires for tire reinforcement. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-020-01623-9>
- Cunningham, P., & Delany, S. J. (2007). k-Nearest neighbour classifiers. *Multi Classification System*, 34(8), 1–17.
- Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, 106263. <https://doi.org/10.1016/j.asoc.2020.106263>
- Diez-Olivan, A., Del Ser, J., Galar, D., & Sierra, B. (2019). Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Information Fusion*, 50, 92–111. <https://doi.org/10.1016/j.inffus.2018.10.005>
- Escobar, C. A., McGovern, M. E., & Morales-Menendez, R. (2021). Quality 4.0: a review of big data challenges in manufacturing. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-021-01765-4>
- Filz, M. A., Gellrich, S., Turetskyy, A., Wessel, J., Herrmann, C., & Thiede, S. (2020). Virtual quality gates in manufacturing systems: Framework, implementation and potential. *Journal of Manufacturing and Materials Processing*. <https://doi.org/10.3390/jmmp4040106>
- Foumani, M., Razeghi, A., & Smith-Miles, K. (2020). Stochastic optimization of two-machine flow shop robotic cells with controllable inspection times: From theory toward practice. *Robotics and Computer-Integrated Manufacturing*, 61, 101822. <https://doi.org/10.1016/j.rcim.2019.101822>
- Foumani, M., & Tavakkoli-Moghaddam, R. (2019). A scalarization-based method for multiple part-type scheduling of two-machine robotic systems with non-destructive testing technologies. *Iranian Journal of Operations Research*, 10(1), 1–17. <https://doi.org/10.29252/iors.10.1.1>
- García, V., Sánchez, J. S., Rodríguez-Picón, L. A., Méndez-González, L. C., de Ochoa-Domínguez, H., & J. . (2019). Using regression models for predicting the product quality in a tubing extrusion process. *Journal of Intelligent Manufacturing*, 30, 2535–2544. <https://doi.org/10.1007/s10845-018-1418-7>
- García-Laencina, P. J., Sancho-Gómez, J. L., & Figueiras-Vidal, A. R. (2010). Pattern classification with missing data: A review. *Neural Computing and Applications*, 19, 263–282. <https://doi.org/10.1007/s00521-009-0295-6>
- Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: The Role of Machine Learning. *IEEE Access*, 5, 20590–20616. <https://doi.org/10.1109/ACCESS.2017.2756872>
- He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In: *IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, pp. 1322–1328. <https://doi.org/10.1109/IJCNN.2008.4633969>
- Hirai, T., & Kano, M. (2015). Adaptive virtual metrology design for semiconductor dry etching process through locally weighted partial least squares. *IEEE Transactions on Semiconductor Manufacturing*, 28(2), 137–144. <https://doi.org/10.1109/TSM.2015.2409299>
- Janitza, S., Strobl, C., & Boulesteix, A. L. (2013). An AUC-based permutation variable importance measure for random forests. *BMC Bioinformatics*, 14, 1. <https://doi.org/10.1186/1471-2105-14-119>
- Kang, P., Lee, H.-J., Cho, S., Kim, D., Park, J., Park, C. K., & Doh, S. (2009). A virtual metrology system for semiconductor manufacturing. *Expert Systems with Application*, 36, 12554–12561. <https://doi.org/10.1016/j.eswa.2009.05.053>
- Kang, S., & Kang, P. (2017). An intelligent virtual metrology system with adaptive update for semiconductor manufacturing. *Journal of Process Control*, 52, 66–74. <https://doi.org/10.1016/j.jprocont.2017.02.002>
- Kao, H. A., Hsieh, Y. S., Chen, C. H., & Lee, J. (2017). Quality prediction modeling for multistage manufacturing based on classification and association rule mining. In: *MATEC web of conferences, the 2nd international conference on precision machinery and manufacturing technology (ICPMMT 2017)*, 123, 00029. <https://doi.org/10.1051/mateconf/201712300029>
- Kim, D., Kang, P., Cho, S., Lee, H. J., & Doh, S. (2012). Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing. *Expert Systems with Applications*, 39, 4075–4083. <https://doi.org/10.1016/j.eswa.2011.09.088>
- Köksal, G., Batmaz, I., & Testik, M. C. (2011). A review of data mining applications for quality improvement in manufacturing industry. *Expert Systems with Applications*, 38(10), 13448–13467. <https://doi.org/10.1016/j.eswa.2011.04.063>
- Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, 31(3), 249–268.
- Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets : A review. *GESTS International Transactions on Computer Science and Engineering*, 30, 25–36.
- Lee, D. H., Yang, J. K., & Kim, K. J. (2020). Multiresponse optimization of a multistage manufacturing process using a patient rule induction method. *Quality and Reliability Engineering International*, 36(6), 1982–2002. <https://doi.org/10.1002/qre.2669>
- Lee, D., Yang, J., Lee, C., & Kim, K. (2019). A data-driven approach to selection of critical process steps in the semiconductor manufacturing process considering missing and imbalanced data. *Journal of Manufacturing Systems*, 52, 146–156. <https://doi.org/10.1016/j.jmsy.2019.07.001>
- Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*, 63(2), 157–170. <https://doi.org/10.1016/j.bushor.2019.10.005>
- Lee, J. (2020). *Industrial AI*. Springer, Singapore. <https://doi.org/10.1007/978-981-15-2144-7>
- Lee, J., Lapira, E., Yang, S., & Kao, A. (2013). Predictive manufacturing system - Trends of next-generation production systems. *IFAC Proceedings*, 46(7), 150–156. <https://doi.org/10.3182/20130522-3-BR-4036.00107>
- Lee, W. J., Xia, K., Denton, N. L., Ribeiro, B., & Sutherland, J. W. (2021). Development of a speed invariant deep learning model with application to condition monitoring of rotating machinery. *Journal of Intelligent Manufacturing*, 32, 393–406. <https://doi.org/10.1007/s10845-020-01578-x>
- Li, F., Wu, J., Dong, F., Lin, J., Sun, G., Chen, H., & Shen, J. (2018). Ensemble machine learning systems for the estimation of steel quality control. *IEEE International Conference on Big Data (big Data)*, 2018, 2245–2252. <https://doi.org/10.1109/BigData.2018.8622583>

- Li, X., Jia, X., Yang, Q., & Lee, J. (2020). Quality analysis in metal additive manufacturing with deep learning. *Journal of Intelligent Manufacturing*, 31, 2003–2017. <https://doi.org/10.1007/s10845-020-01549-2>
- Lieber, D., Stolpe, M., Konrad, B., Deuse, J., & Morik, K. (2013). Quality prediction in interlinked manufacturing processes based on supervised & unsupervised machine learning. *Procedia CIRP*, 7, 193–198. <https://doi.org/10.1016/j.procir.2013.05.033>
- Lipton, Z. C., & Steinhardt, J. (2019). Troubling trends in machine-learning scholarship. *Queue*, 17(1), 1–15. <https://doi.org/10.1145/3317287.3328534>
- Malaca, P., Rocha, L. F., Gomes, D., Silva, J., & Veiga, G. (2019). Online inspection system based on machine learning techniques: real case study of fabric textures classification for the automotive industry. *Journal of Intelligent Manufacturing*, 30, 351–361. <https://doi.org/10.1007/s10845-016-1254-6>
- McCann, M., & Johnston, A. (2008). SECOM dataset UCI machine learning repository. <https://archive.ics.uci.edu/ml/datasets/SECOM>
- Melhem, M., Ananou, B., Djeziri, M., Ouladsine, M., & Pinaton, J. (2015). Prediction of the wafer quality with respect to the production equipments data. *IFAC-PapersOnLine*, 48, 78–84.
- Mineo, C., Pierce, S. G., Nicholson, P. I., & Cooper, I. (2016). Robotic path planning for non-destructive testing - A custom MATLAB toolbox approach. *Robotics and Computer-Integrated Manufacturing*, 37, 1–12. <https://doi.org/10.1016/j.rcim.2015.05.003>
- Moldovan, D., Cioara, T., Anghel, I., & Salomie, I. (2017). Machine learning for sensor-based manufacturing processes. In: *2017 13th IEEE international conference on intelligent computer communication and processing (ICCP)*, pp. 147–54. <https://doi.org/10.1109/ICCP.2017.8116997>
- Munirathinam, S., & Ramadoss, B. (2016). Predictive models for equipment fault detection in the semiconductor manufacturing process. *International Journal of Engineering and Technology*, 8(4), 273–285. <https://doi.org/10.7763/ijet.2016.v8.898>
- Nakagawa, S. (2015). Missing data: mechanisms, methods, and messages. *Ecological Statistics: Contemporary Theory and Application*. <https://doi.org/10.1093/acprof:oso/9780199672547.003.0005>
- Nti, I. K., Adekoya, A. F., Weyori, B. A., & Nyarko-Boateng, O. (2021). Applications of artificial intelligence in engineering and manufacturing: a systematic review. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-021-01771-6>
- Ou, X., Huang, J., Chang, Q., Hucker, S., & Lovasz, J. G. (2020). First time quality diagnostics and improvement through data analysis: A study of a crankshaft line. *Procedia Manufacturing*, 49, 2–8. <https://doi.org/10.1016/j.promfg.2020.06.003>
- Peres, R. S., Barata, J., Leitao, P., & Garcia, G. (2019). Multistage quality control using machine learning in the automotive industry. *IEEE Access*, 7, 79908–79916. <https://doi.org/10.1109/ACCESS.2019.2923405>
- Quatrini, E., Costantino, F., Di Gravio, G., & Patriarca, R. (2020). Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities. *Journal of Manufacturing Systems*, 56, 117–132. <https://doi.org/10.1016/j.jmsy.2020.05.013>
- Raschka, S. (2018). Model evaluation, model selection, and algorithm selection in machine learning. arXiv: 1811.12808
- Rezaei-Malek, M., Mohammadi, M., Dantan, J. Y., Siadat, A., & Tavakkoli-Moghaddam, R. (2019). A review on optimisation of part quality inspection planning in a multi-stage manufacturing system. *International Journal of Production Research*, 57, 4880–4897. <https://doi.org/10.1080/00207543.2018.1464231>
- RStudio Team. (2020). RStudio: Integrated Development Environment for R
- Salem, M., Taheri, S., & Yuan, J. S. (2018). An experimental evaluation of fault diagnosis from imbalanced and incomplete data for smart semiconductor manufacturing. *Big Data and Cognitive Computing*, 2(4), 30. <https://doi.org/10.3390/bdcc2040030>
- Sarkar, B., & Saren, S. (2016). Product inspection policy for an imperfect production system with inspection errors and warranty cost. *European Journal of Operational Research*, 248(1), 263–271. <https://doi.org/10.1016/j.ejor.2015.06.021>
- Schmitt, J., Böning, J., Borggräfe, T., Beitinger, G., & Deuse, J. (2020). Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing. *Advanced Engineering Informatics*, 45, 101101. <https://doi.org/10.1016/j.aei.2020.101101>
- Tadist, K., Najah, S., Nikolov, N. S., Mrabti, F., & Zahi, A. (2019). Feature selection methods and genomic big data: a systematic review. *Journal of Big Data*. <https://doi.org/10.1186/s40537-019-0241-0>
- Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*, 48, 157–169. <https://doi.org/10.1016/j.jmsy.2018.01.006>
- Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*, 17(1), 168–192. <https://doi.org/10.1016/j.aci.2018.08.003>
- Thiede, S., Turetskyy, A., Kwade, A., Kara, S., & Herrmann, C. (2019). Data mining in battery production chains towards multi-criterial quality prediction. *CIRP Annals*, 68, 463–466. <https://doi.org/10.1016/j.cirp.2019.04.066>
- Tsai, C. F., Hsu, Y. F., Lin, C. Y., & Lin, W. Y. (2009). Intrusion detection by machine learning: A review. *Expert Systems with Applications*, 36, 11994–12000. <https://doi.org/10.1016/j.eswa.2009.05.029>
- Turetskyy, A., Thiede, S., Thomitzek, M., von Drachenfels, N., Pape, T., & Herrmann, C. (2020). Toward data-driven applications in lithium-ion battery cell manufacturing. *Energy Technology*, 8(2), 1900136. <https://doi.org/10.1002/ente.201900136>
- Venkatesh, B., & Anuradha, J. (2019). A review of feature selection and its methods. *Cybernetics and Information Technologies*, 19(1), 3–26. <https://doi.org/10.2478/CAIT-2019-0001>
- Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., & Wrobel, S. (2019). A review of machine learning for the optimization of production processes. *The International Journal of Advanced Manufacturing Technology*, 104, 1889–1902. <https://doi.org/10.1007/s00170-019-03988-5>
- Wuest, T., Irgens, C., & Thoben, K. D. (2014). An approach to monitoring quality in manufacturing using supervised machine learning on product state data. *Journal of Intelligent Manufacturing*, 25, 1167–1180. <https://doi.org/10.1007/s10845-013-0761-y>
- Xu, S., Lu, B., Baldea, M., Edgar, T. F., Wojsznis, W., Blevins, T., & Nixon, M. (2015). Data cleaning in the process industries. *Reviews in Chemical Engineering*, 31(5), 453–490. <https://doi.org/10.1515/revce-2015-0022>
- Yin, X., He, Z., Niu, Z., & Li, Z. (2018). A hybrid intelligent optimization approach to improving quality for serial multistage and multi-response coal preparation production systems. *Journal of Manufacturing Systems*, 47, 199–216. <https://doi.org/10.1016/j.jmsy.2018.05.006>
- Yugma, C., Blue, J., Dauzère-Pères, S., & Obeid, A. (2015). Integration of scheduling and advanced process control in semiconductor manufacturing: review and outlook. *Journal of Scheduling*, 18, 195–205. <https://doi.org/10.1007/s10951-014-0381-1>
- Zhai, J., Xu, X., Xie, C., & Luo, M. (2004). Fuzzy control for manufacturing quality based on variable precision rough set. In: *Fifth world congress on intelligent control and automation (IEEE Cat. No.04EX788)*, Vol. 3, pp. 2347–2351. <https://doi.org/10.1109/wcica.2004.1342013>
- Zhang, W., Li, X., & Ding, Q. (2019). Deep residual learning-based fault diagnosis method for rotating machinery. *ISA Transactions*, 95, 295–305. <https://doi.org/10.1016/j.isatra.2018.12.025>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.