Decision Analytics Journal 3 (2022) 100075

Contents lists available at ScienceDirect

Decision Analytics Journal

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

A data analytics model for improving process control in flexible manufacturing cells

Giovanna Martínez-Arellano [a,*], ThuBa Nguyen [b], Chris Hinton [c], Svetan Ratchev [a]

[a] Institute for Advanced Manufacturing, Jubilee Campus, University of Nottingham, Nottingham, NG72GX, UK
[b] School of Aerospace, Transport and Manufacturing, Cranfield University, UK
[c] Hams Hall Plant, BMW Group, Birmingham, UK

ARTICLE INFO

Keywords:
Flexible manufacturing systems
Statistical process control
Data analytics
Data visualisation

ABSTRACT

With the need of more responsive and resilient manufacturing processes for high value, customised products, Flexible Manufacturing Systems (FMS) remain a very relevant manufacturing approach. Due to their complexity, quality monitoring in these types of systems can be very difficult, particularly in those scenarios where the monitoring cannot be fully automated due to functional, safety and legal characteristics. In these scenarios, quality practitioners concentrate on monitoring the most critical processes and leaving out the inspection of those that are still meeting quality requirements but showing signs of future failure. In this paper we introduce a methodology based on data analytics that simplifies the monitoring process for the operator, allowing the practitioner to concentrate on the relevant issues, anticipate out of control processes and take action. By identifying a *reference model* or best performing machine, and the occurring patterns in the quality data, the presented approach identifies the adjustable processes that are still in control, allowing the practitioner to decide if any changes in the machine's settings are needed (tool replacement, repositioning the axis, etc.). An initial deployment of the tool at BMW Plant Hams Hall to monitor a focussed set of part types and features has shown a reduction in scrap of 97% throughout 2020 in relation to the monitored features compared to the previous year. This in the long run will reduce reaction time in following quality control procedure, reduce significant scrap costs and ultimately reduce the need for measurements and enable more output in terms of volume capacity.

1. Introduction

Flexibility in manufacturing systems is becoming more and more necessary as domestic and global markets rapidly change, technology evolves and demand towards more customisable products increases. Flexibility in manufacturing allows companies to be able to deal with fast changing product types, production volumes, assembly variation and process sequence while keeping production processes time and cost effective [1,2]. FMS achieve these goals with programmable automation through the use of a number of systems such as CNC machine tools, interconnected by automated material handling and storage systems, inspection stations and gauging systems [3].

The reliability of the automated equipment in FMS is a critical contributor to system performance, and so the capability of performing correct and rapid fault diagnosis and process variation identification through Statistical Process Control (SPC) is essential [4]. SPC and the use of control charts can detect process variations and identify possible machine breakdowns and out-of-control processes that can affect the availability of resources. However, it has been found that traditional SPC approaches are usually not appropriate for production paradigms such as flexible and re-configurable systems, where the production is small or there is a high variety of mixed products. Short production runs pose several challenges. First, the lack of available data to estimate reliable process parameters due to smaller production runs. This is a typical scenario in Just-in-time (JIT) systems, where low levels of inventory are kept, and during start-up of a process or initiation of a new process, where there is an insufficient number of subgroups of measurements under different conditions available [5]. Second, there is an increasing risk of false acceptance in SPC because of measurement errors [6]. Several possible sources of error are equipment accuracy, operator mistakes, environmental factors and random noise. Measurement errors can lead to unnecessary process adjustments and loss of confidence in SPC [7]. Third, the production of customised products and services are characterised by complex designs and processes, requiring more flexible quality management practices. SPC in flexible environments requires

* Corresponding author.
E-mail addresses: Giovanna.martinezarellano@nottingham.ac.uk (G. Martínez-Arellano), ThuBa.Nguyen@cranfield.ac.uk (T. Nguyen), Chris.hinton@bmwgroup.com (C. Hinton), svetan.ratchev@nottingham.ac.uk (S. Ratchev).

https://doi.org/10.1016/j.dajour.2022.100075
Received 27 December 2021; Received in revised form 23 May 2022; Accepted 25 May 2022
Available online 2 June 2022

more control charts to monitor quality characteristics and product specifications compared to more static approaches [8]. The data generated by analysing and monitoring individual process, product family and test stations can be extremely large and hard to manage [9]. Manufacturing companies still rely on operator and quality practitioner experience in manually analysing the information. There are product characteristics within manufacturing which need particular attention because of their functional, safety and legal characteristics requirements [10]. These types of features often carry tight tolerances or parameters which are particularly sensitive to variation and change in the production systems. This is why employees who are involved in the collection and analysis of the data need to understand statistical concepts and how they can be interpreted to process control [11].

In a flexible manufacturing system, where the number of machines, component variants and product characteristics to control is so high, traditional control charts used for quality control are not effective anymore due to "curse of dimensionality" [12]. The operator will prioritise the analysis and reaction method of those processes which tolerance limits are breached, disregarding other processes that might not be running at the expected standards. Here is where data analytics and Machine Learning (ML) based frameworks implementing self-awareness, self-diagnosis, self-prognosis and self-healing promise to be beneficial. By having a system that can continuously monitor its state and provide any relevant change to the practitioner at the time needed, will allow the expert to concentrate in more relevant tasks, even if the system is not completely automated in the healing aspect. However, despite the development of new process control methods based on technologies such as ML, these normally solve some but not all the issues [8]. With the overwhelming number of machine learning and data analytics techniques available to the practitioner and the lack of experience to determine the necessary pre-processing steps for these to work effectively, new methods turn out very difficult to implement by the non-expert, and so there is very little evidence of these new techniques being successful in industry.

This paper takes a different angle from state-of-the art approaches. Due to the varying availability of measurements across different key characteristics, part types and machines, which makes most ML methods unsuitable, this paper proposes to semi-automate and support manual process control through the use of existing SPC practices that ensure quality standards are met (e.g. IATF16949), but simplifying and prioritising the decisions the operator needs to make and ensuring focus is paid to characteristics that are showing early signs of problems. This is achieved through a methodology based on identifying the best performing machine and the set of adjustable machines, using the former as a reference to adjust other machines. The proposed methodology uses classical SPC together with advanced data analytics algorithms to automate the analysis of quality data in near real-time, allowing the practitioner to anticipate problems before processes go out of control. The approach is used and evaluated in a flexible manufacturing environment at BMW, where the number of machines, component variants and features per component are high, demonstrating the benefit of the tool compared to the typical monitoring methods used in such environment. By combining the use of traditional SPC with state-of-the art non-parametric techniques, current challenges of traditional SPC methods which are still widely used in manufacturing environments, are addressed and demonstrated. The paper makes the following contributions:

- It deals with errors in measurements (outliers), which have a major impact in statistical features used in typical SPC approaches. It is typically down to the expertise of the quality practitioner to identify these measurements and leave them out of the analysis.
- Proposes a reference model and alert level methodology based on well known quality standards, which provides an effective automated characterisation of the performance of a machine when dealing with a large number of features but low volumes of data. This data challenge is nowadays more common with the production of more customised, high value products and batch-of-one demands and poses challenges to advanced machine learning models that depend on large amounts of data being available.
- It takes advantage of the lead model to perform not only process improvement in a single machine, but across multiple machines, implementing a *group awareness* figure which will be essential in future process improvement solutions.
- It demonstrates in a real industrial scenario a successful implementation of a human in the loop monitoring–diagnosis–prognosis framework using the proposed methodology.

The rest of the paper is organised as follows: Section 2 presents work related to recent developments of SPC techniques particularly for FMS. The methodology is then presented in Section 3, followed by a detailed description of the case study at BMW and a discussion of results in Section 4. Finally, Section 5 presents the conclusions and future work.

2. Related work

Although flexible manufacturing is not a new paradigm, there has been growing interest in recent years due to factors including the growing demand for customised products, shorter product life cycles and environmental impact [13,14]. Compared to traditional transfer lines that enable high volume but less flexibility in their design, FMS allow for reaction to volume fluctuations, mixed batch production and multiple product types within the same product family [15]. This flexibility involves having machine capabilities that are themselves flexible (such as CNC machines) and hence can produce multiple product variants, but at the cost of making production lines more complex and difficult to monitor. Like in any manufacturing paradigm, SPC plays an important part in FMS to ensure that manufacturing processes operate in their in-control state. However, it has been found that traditional SPC methods are not appropriate for situations of small lots or where a high variety of products exist [5]. Although flexible production systems may manufacture large volumes, the production in these types of environments is intermittent because the change to other product variants is easy. This intermittent aspect makes traditional SPC methods particularly unsuitable. In addition, whilst being highly automated, the balance required between quality control interruption and intelligence driven solutions needs to be addressed. A typical process within a production line in an FMS would require the inspection of tens of thousands of control charts, making it possible for the operator to miss processes or features which have been affected or shifted but still fall within acceptable limits.

2.1. SPC control charts

Traditional control charts such as Shewhart, Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA) assume that the values of a Process Mean and Variance are known *a priori* at the start of the process monitoring or that the data for estimating the process parameters is available during a production run [16]. Traditional SPC is based on the assumption that the data is Independent and Identically Normally Distributed (IIND). However, in environments such as an FMS where there is a variety of mixed products processed in the same production line, these fundamental assumptions are not met. As different measured part features may have different means and variances, the IIND condition is not satisfied.

There are many methods that have been developed to address the limitations of classical SPC methods in FMS. Hillier developed methods to adjust control limits used in X-bar and R control charts [17]. Quesenberry proposed Q charts, which can detect changes in the mean and variance in short production runs [18]. These charts have been further studied and enhanced in other works [19,20], as it has been found that the method is sensitive to early shifts of the mean, or when the variance is unknown. T-charts are other new type of control charts based on the calculation of the T-statistic and were first proposed by

Zhang et al. [21]. A T-chart does not require a preliminary estimation of the in-control process Standard Deviation (SD), making it useful when monitoring at the startup of a process. However, T-charts rely on setting up the control mean at the beginning of the production run. Further developments of T-charts have been done by [22,23]. Many other approaches have dealt with the normality and non independent assumption validations as well as the small sample size such the ARMA chart [24], ML-chart [25], ACUSUM charts [26].

2.2. Change-point SPC

In general, control charts allow to determine if a process is in or out of control. However, in some circumstances, by the time the process has reached the control limit and an alert has been raised, it might be already too late to avoid or fix the related issue. An early detection of a trend or pattern in the measurements before reaching the control limit might allow the operator to identify and react on time to a possible future issue. To complement charts, methods to detect shifts in short-run situations such as the change-point method have been proposed [27]. The basic change-point approach, as introduced by Hawkins, is based on the assumption that process readings can be modelled by two normal distributions. Process readings follow an initial distribution until certain point τ, the *change-point*, at which point they switch to another normal distribution, differing in the mean, variance or both [28]. Methods to monitor the process mean and process variance in univariate and multivariate applications using the change-point method have been proposed in [29,30]. In short-run environments where data is scarce, outliers can have a big impact on detecting true change-points. Methods with particular focus on outliers have also been proposed [31,32]. To deal with data that is not normally distributed, non-parametric solutions have been proposed as well [33,34]. Although a considerable amount of work has been done on these methods, there is still work to be done, particularly in non-parametric techniques [35]. Jones-Farmer et al. particularly stress that for more quality-control practitioners to use these methods on real scenarios, there needs to be more development of easy-to-use computer software that can make these complex methods accessible to them [35].

2.3. Control chart pattern recognition

In addition to change-point based strategies, there are other methods based on ML that have been applied for detecting patterns in process quality data. Different detected patterns may be an indication of different problems. While trend patterns may indicate a slow change, cyclic patterns might be related to periodic variation in the process such as power supply [36]. In the early application of control charts, it was necessary to manually determine whether or not there was an abnormality, this being more difficult when the process was still in control. To this, many pattern detection methods based on supplementary rules where proposed [37,38]. Wang et al. defined 30 possible control chart patterns, categorised as 8 basic patterns and 22 combinations of the basic ones [37]. However, dealing with a large number of rules may be difficult when doing real-time monitoring. To address this, recent work in applying ML for pattern recognition has been done with the aim to automate this process [36,39]. An exhaustive review of techniques can be found in [40].

Control chart pattern recognition methods have been incorporated with other SPC methods in an attempt to provide more complete process monitoring tools (i.e. online monitoring, diagnosis and prescription) that could support the quality practitioner to take action in real time. Guh developed an on-line SPC system that incorporates an Artificial Neural Network (ANN) to perform pattern recognition together with an expert system that can interpret or relate the patterns recognised with possible issues and propose solutions to the quality practitioner [41]. Other online process monitoring and diagnosis approaches that combine ML techniques have been reported [42,43].

Root cause analysis and diagnosis has been a main focus of the research in process monitoring in recent years. However, there is another dimension to process monitoring that is still to be reached: *Prognosis* i.e. determining the likely development and change of a problem. Current trends in industrial process monitoring are moving towards this direction [44,45]. Despite the amount of work published in ML-based monitoring and diagnostic methods, most of these are tested in a simulated scenario, only a few industrial case studies for these tools have been reported [46]. In addition, more flexibility is still needed so that these types of approaches can be used by practitioners. Most ML applications are still inaccessible to most practitioners due to their complexity and required sills needed for their implementation and continued use.

2.4. Computerised tools in real industry scenarios

Along with the development of new methods for improving process control, there have been efforts for implementing these in real scenarios which involve the development of computational tools for SPC implementation and data visualisation. Azadeh and Zeynali developed a framework of integrated quality control that features a Quality Information System, Statistical Quality Control, and SPC. The system was tested for a 5-year period in a large industrial machinery manufacturer, demonstrating the benefits of the computerised system [47]. Huang et al. performed a case study in a Taiwanese LCD manufacturer where they study the key factors that can make an SPC system successful by developing an effective performance evaluation model. The authors claim that despite the success of large automotive companies implementing the SPC manual in QS-9000 (quality standard which has been around since 1994), most organisations are still learning about the most effective ways to introduce, develop and implement SPC. One of the key factors that influenced the success of the implementation of SPC in a business was adequate training and support for the users of the system through either internal education or external experts and advisors [48]. Guerra et al. developed an SPC software tool that was used to automate the quality control in the final inspection process of a production line in an automotive company. The authors stressed the lack to published work related to case studies in industry and how industry is dealing with the implementation of computerised SPC systems [49]. Liang et al. propose a fog and deep learning-based prognosis system which is validated in a UK machining company. The authors stressed the fact that most cloud and deep learning based approaches for prognosis are ineffective to meet the requirements of practical manufacturing processes [46]. More recently, Schmitt et al. developed a predictive model-based quality inspection using ML and edge cloud computing which was validated with an industrial case study [50]. However, the development of such ML solution requires of complex steps such as data processing, model selection and training, which makes it still inaccessible to most practitioners [51].

2.5. Big data analytics and frameworks for Industry 4.0

With the advances in Information Technology (IT) and increased computing power, data analytics at both small and large scales are now possible. There is a vast amount of work done in the development and implementation of advanced data analytics methods for condition monitoring [52], fault diagnosis [53], anomaly detection [54] among other process industry applications; and these developments are often considered part of Industry 4.0. In fact, these monitoring and diagnosis abilities incorporated into the production system itself and/or its machines, referred in the literature as self-awareness and self-diagnosis, and the recognition the state of other systems (group awareness) is expected to be a key characteristic of Industry 4.0 [55], where the human expert will still play a key role in the continuous process improvement loop as proposed in the framework by Cohen and Singer [56]. However, despite the vast amount of research on Industry

4.0 architecture technologies and applications, industries are still not confident enough on the implementation of these technologies due to the unclear benefits as well as the lack of implementation details and the investment needed [57]. The challenges that still remain in industry are:

- Process complexity - monitoring a large amount of process parameters/quality features to identify issues as they rely heavily on the quality practitioner's expertise.
- Data size - Large data volumes in terms of number of process variations/parameters but lack of measurements of one specific process variation in a FMS environment.
- Technical expertise - there is still scepticism from industry in regards to ML applications, and a lack of practitioners with the necessary skills.

This work attempts to address these challenges through the development and implementation of a process monitoring methodology and visualisation tool. The methodology allows to identify normal and out-of-control process behaviour and goes a bit further by allowing the practitioner to anticipate potential issues.

3. Methodology

By exploiting the current advances in data analytics, this work presents a methodology that incorporates classic and advanced SPC methods together with data analytics algorithms to perform the identification of potential problems in processes before these reach the specified control limits. This is performed by identifying, -in an online manner-, the best performing process (e.g. machine) with as few as 5 measurements and using this to identify other processes that can be potentially improved as their performance moves away from the best case. As it is shown in Fig. 1, there are four major steps performed in the aforementioned methodology: *data pre-processing, reference model identification, machine characterisation* and *visualisation*. These are described in more detail in the following subsections.

3.1. Data pre-processing

FMS are characterised for the complexity and variation of its processes. The methodology described in this work is intended for FMS environments, where the number of process variations and features within one process variation is large. Particularly, this work is applied in an automotive industrial environment with the following characteristics:

- There is a number of component variations (i.e. part types) depending on demand
- Each part type has hundreds of product characteristics (i.e. features) to monitor across multiple machines.
- One machine is capable of producing the same features of multiple components variants to achieve high volume production.

These system factors determine the data complexity. The number of features (e.g. geometrical features measured with a CMM) for each part type variant can be very large, but at the same time, the number of continuous measurements for each machine for a particular feature/part variant can be small. Data pre-processing must occur due to the sensitivity of the measurement devices, machining and measurement processes and subsequent data outputs. Data from industrial multimode processes with different operating conditions and transitions are affected by noise, outliers and missing data that affect the precision of classic monitoring approaches [58]. This requires engineers within the business to have a level of proficiency in data pre-processing by adapting their skills around query languages and rule-based methodology. The benefits of approaching these automated cleaning steps simplify the process [9] and put the focus back to the engineering decisions which need to be made. Before performing any statistical analysis related to the performance of a machine using product features, the data is pre-processed by eliminating any duplicate measurements and outliers that can be related to errors in the CMM measurements. As mentioned previously, outliers can have a large effect on the estimation of process control statistics, particularly when data is scarce. Most common outlier detection methods are based on the SD over the mean. However, this is fundamentally problematic, as the indicator itself is altered by the presence of outliers. The absolute deviation from the Median Absolute Deviation (MAD), on the other hand, while still being a measure of central tendency, is insensitive to the presence of outliers and immune to sample size, making it a better indicator [59]. Calculating the MAD involves finding the median of absolute deviations from the median. It is defined as follows:

$$MAD = bM_i(|x_i - M_j(x_j)|) \tag{1}$$

where the x_j is the n original observations, M_i is the median of the series and b is a constant, usually $b = 1.4826$, linked to the assumption of normality of the data. The constant b ensures that for large samples, the MAD remains a consistent estimator of the SD of the population. The decision criteria is then defined as:

$$M - t * MAD < x_i < M + t * MAD \tag{2}$$

where t is the threshold selected [60]. In this methodology, MAD is applied on all quality measurements per machine and per feature, and across multiple part types on a given time frame. An example of outliers detected for a feature F in a set of machines can be seen in Fig. 2.

3.2. Statistical analysis to determine reference model and adjustable machines

During monitoring, reference models can be an effective way to detect process deviations [61]. However, generating these reference models requires of expert-knowledge or relies on large amounts of data being available [58]. In this work, a reference model or best machine is proposed as a first step to start characterising the performance of machines, however, it does not rely solely on this concept to identify issues, as it will be explained in the next subsections. The best performing machine in a particular period of time (determined by the user but at least 5 measurements required) is identified by calculating for each feature measurement its deviation from the nominal or target value, this could be the real deviation or the percentage. The deviation of a measurement m is calculated as follows:

First, the nominal value related to each measurement is determined. This value is zero in the case of unilateral tolerances, otherwise the nominal value is:

$$nominalValue = uCL - lCL \tag{3}$$

where uCL stands for upper control limit and lCL for lower control limit. From this nominal value, then a deviation from target can be calculated:

$$targetDeviation_m = |nominalValue - m| \tag{4}$$

For the percentage deviation, if it is the case of an unilateral tolerance then:

$$percentageDeviation_m = \frac{m}{uCL} \times 100 \tag{5}$$

If the lower control limit of the feature is different from zero:

$$percentageDeviation_m = \frac{targetDeviation_m}{nominalValue \times \frac{1}{2}} \times 100 \tag{6}$$

The percentage deviation is useful when using multiple features for the best machine calculation when these might have different tolerances. In a second step, the target deviation is used to calculate a set of statistical features per feature and machine. These include:

- Mean, SD and Variance of the feature measurement m.

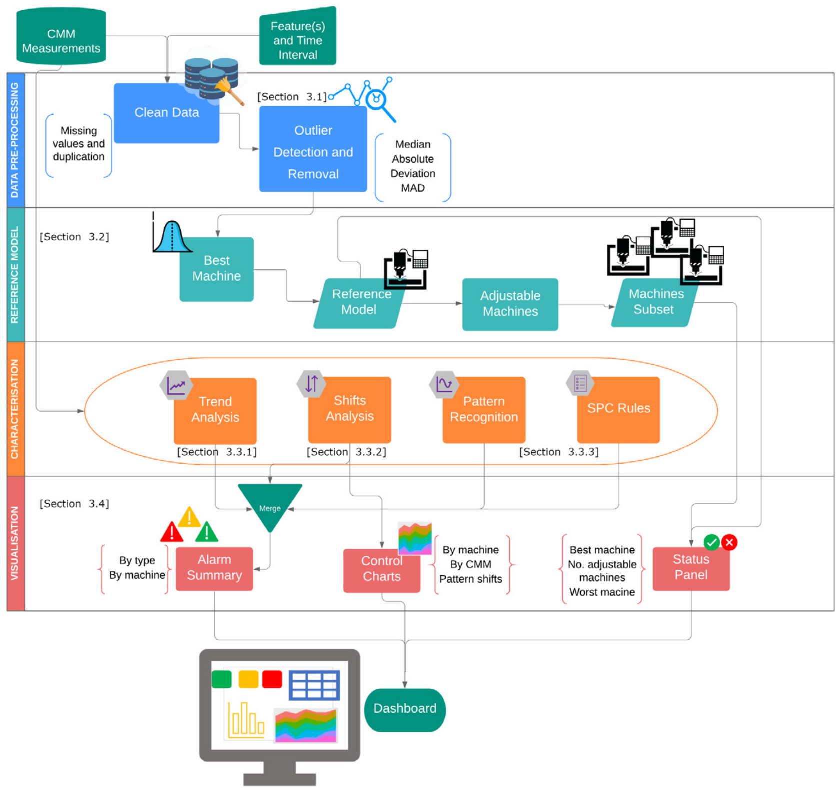

Fig. 1. Methodology used to automatically identify in and out-of-control processes as well as anticipate potential problems occurring at different machines in a flexible manufacturing line.

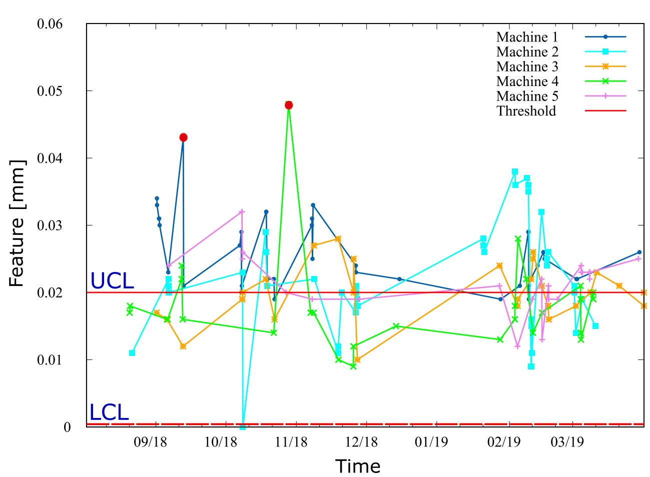

Fig. 2. Outliers (red dots) identified on CMM measurements of a particular feature F corresponding to 5 machines. Although the algorithm picks up an outlier in machine A_7_1, this measurement is not considered an outlier in this particular case as the target value for this feature is zero. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

- Minimum and maximum measurement.
- Mean and SD of the deviation of the measurement from the target value.
- Minimum and maximum deviation from target.
- Number of measurements out of threshold

These statistical features will be used for the best machine estimation as well as for constructing the visualisation dashboard which provides not only an estimated best machine but also statistical information on other machines that could be performing similarly.

The *best machine* is then defined, in a first instance, as the machine that, for a feature or set of features, has at least five measurements for the time period currently being looked at, does not have any of those measurements out of threshold and has the smallest Mean deviation. In the same way, the *worst machine* is a machine with at least five available measurements which has the highest Mean deviation. From here, a set of *adjustable machines* are defined as those machines that have a higher Mean deviation compared to the best machine but their measurements fall within 3 SDs and within the upper and lower control limits. This does not include those machines that have a lower Mean deviation compared to the best machine but less than five measurements available. Fig. 3 shows an example of the best machine found for a feature F as well as one of the identified adjustable machines. The blue doted lines show the 3SD of the machine identified as the best one. As it can be observed, machine A_3_2 is just slightly worse than A_4_1.

The *best machine*, *worst machine* and set of *adjustable machines* as defined previously will be the first indicator of performance of the process for a particular feature or set of features of one or more part types. If all machines apart from the best machine are adjustable, then the process is in-control and can be potentially improved according to the parameters of the best machine. Whilst machine adjustments and changes happen independently, each machine will behave differently but still have the same technology and approach to setting. If a particular machine needs adjusting to the point where the technique is not completely clear, then a reference model is useful to help guide the engineers to knowing“what good looks like” and using that as a base guide, by comparing for example tool wear, parameter setting, spindle vibration, among other data available to the operator. To support this problem identification process, it is necessary to complement this information by characterising further through patterns and rules presented in the following subsections. If no best machine is found, is either because there are still not enough measurements to determine the performance or because all machines are showing measurements out of threshold. In any case, as it will be discussed in the industrial

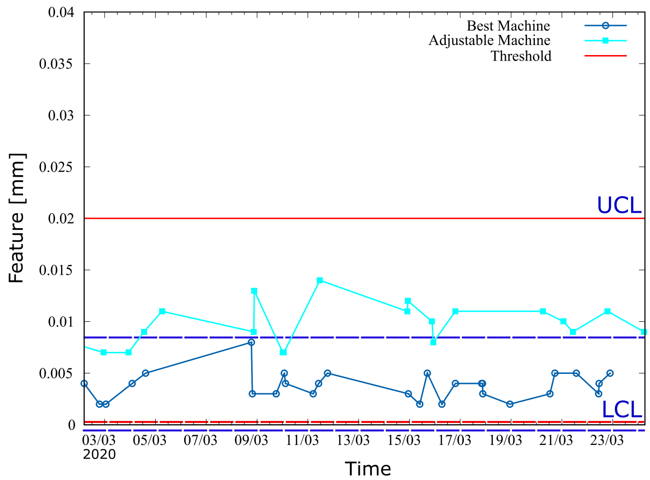

Fig. 3. Example where machine A_4_1 is identified as the best machine (closer to the target value zero) and A_3_2 is identified as one of the adjustable machines.

case, through the continuous logging of adjustments as a response to particular machine performance, it is possible to start building a knowledge base that facilitates future machine adjustment even when a best model is not identified.

3.3. Pattern analysis for the characterisation of machine performance

The way in which each machine is further characterised is through the automated identification of different patterns and rules. In classical SPC, control chart patterns or behaviours typically relate to different process issues. Based on the basic control chart patterns that have been defined in the literature [40], in this work trends, shifts and rules are applied and results integrated to automate the characterisation of a machine performance. While some patterns may indicate a re-occurring problem, some may indicate an improvement. For example, an increasing/decreasing trend on the measurements approaching the target value would indicate an improvement. Each pattern will have an associated alert or level of severity, either green, yellow, or red, depending on the nature of the change. Once all machines are checked for patterns and rules, a summary of alerts is generated. By integrating the resulting information of all patterns found and relating them to a level of severity, it is possible to identify scenarios where a problem is beginning to develop, particularly when similar patterns are found across multiple features. This will be further discussed in the visualisation section.

3.3.1. Increasing/decreasing trends

A trend is present if a sequence of measurements exhibits steady upward or downward decline over its whole length. To identify a trend, the Cox-Stuart test is performed. This is a non-parametric sign test for detecting trends in independent, time-ordered data [62]. Although this test is not as powerful as the Mann–Kendall test [63], the computational effort of this test is lower (increases linearly with the sequence size). The main steps to perform the Cox-Stuart test are shown in Algorithm 1.

Once the trends for each feature/machine are determined, an alert associated with each trend will be recorded. A detected trend that moves away from the nominal will raise a red alert, whilst a trend towards the nominal will raise an amber alert.

3.3.2. Up/down shifts

Detecting multiple change points in a sequence of values can be computationally expensive. A brute force exact approach would take $O(Qn^2)$ calculations for a sequence of values of size n, where Q is the maximum number of change points. The objective is to find the set of change points τ that minimise the equation:

$$\sum_{i=1}^{m+1}[C(y_{(\tau_{i-1}+1):\tau_i})] + \beta f(m) \quad (7)$$

where C is a cost function (typically the negative log-likelihood) for a segment, m is the number of change points (m+1 segments) and $\beta f(m)$ if a penalty to avoid over fitting. It is still an open question how to define a maximum number of change points in a sequence whilst avoiding over fitting.

To check for sudden shifts on the measurements, the Pruned Exact Linear Time (PELT) change point algorithm is used here, as it uses a non-parametric cost function based on the empirical distribution [64]. The PELT algorithm is an exact algorithm which is computationally less expensive compared to algorithms such as Binary Segmentation (linear in the best case scenario) by using dynamic programming and pruning. The PELT algorithm is based on the idea that the optimal number of change points of a segment $y_{1:s}$ ($F(s)$) can be expressed in terms of the optimal number of change points for $y_{1:t}$ where $t < s$. This allows to define the optimal number of change points for $y_{1:s}$ recursively in terms of the minimal cost for $y_{1:t}$. To maintain a linear execution time on n, this algorithm applies pruning by removing those values of τ that can never be minima from the minimisation performed at every iteration. The main assumption of the PELT algorithm is that the number of change points increases linearly as the data set grows, in other words, the change points are spread through the data rather than confined to one portion. The steps of this method are presented in Algorithm 2. Fig. 4 shows the detected shifts on a subset of measurements for a feature and machine. The change points (vertical lines) allow to identify state borders; this means points in times that indicate the end of a particular distribution before changing to a new distribution.

Once all change points are found, for each change point detected in a set of measurements if it represents an increase or decrease in 25% of the tolerance from its previous change point, it is highlighted with a red alert. In addition, for any change point detected in the last 5 observations of a set of measurements, a green alert is raised for those approaching the target value and a red one for those moving away from the target.

input : A set of independent observations of the form $(y_1, y_2, ..., y_n)$ where $y_i \in \Re$
output: A p-value to determine the significance of the inc/dec trend

1. Divide the input data into two vectors $firstHalf = y_{1:(n+1)/2}$ and $secondHalf = y_{(n+1)/2:n}$;
2. Substract value by value the two vectors $dif = firstHalf - secondHalf$;
3. For each difference dif_i obtain the sign ;
4. Calculate the number of positive and negative signs pos and neg ;

if $pos \leq neg$ **then**
  5. Calculate the probability to obtain $x = pos$ number of successes on N=n experiments $prob = pbnom(pos, n, 0.5)$;
  6. If $prob \leq 0.05$, there is an increasing trend, else there is no significant trend ;
else
  7. Calculate the probability to obtain $x = neg$ number of successes on N=n experiments : $prob = pbnom(neg, n, 0.5)$;
  8. If $prob \leq 0.05$, there is a decreasing trend, else there is no significant trend ;
end

Algorithm 1: COX-Stuart Test algorithm

input : A set of data of the form $(y_1, y_2, \ldots, y_n)$
where $y_i \in \Re$
A measure of fit $C(.)$ dependant on the data
A penalty constant β which does not depend
on the number or location of the change points
A constant K that satisfies
$C(y_{(t+1):s}) + C(y_{(s+1):T}) + K \leq C(y_{(t+1):T})$
output: The change points recorded in $cp(n)$
Initialise: let n = length of data and set $F(0) = -\beta$, $cp(0) = NULL$, $R_1 = 0$;
for $\tau^* = 1 \ldots n$ **do**
  1. Calculate $F(\tau^*) = \min_{\tau \in R_{\tau^*}} [F(\tau) + C(y_{(\tau+1):\tau^*}) + \beta]$;
  2. Let $\tau^1 = arg \left\{ \min_{\tau \in R_{\tau^*}} [F(\tau) + C(y_{(\tau+1):\tau^*}) + \beta] \right\}$;
  3. Set $cp(\tau^*) = [cp(\tau^1), \tau^1]$;
  4. Set $R_{\tau^*+1} = \left\{ \tau^* \cap \left\{ \tau \in R_{\tau^*} : F(\tau) + C(y_{(\tau+1):\tau^*}) + K < F(\tau^*) \right\} \right\}$;
end

Algorithm 2: PELT Method

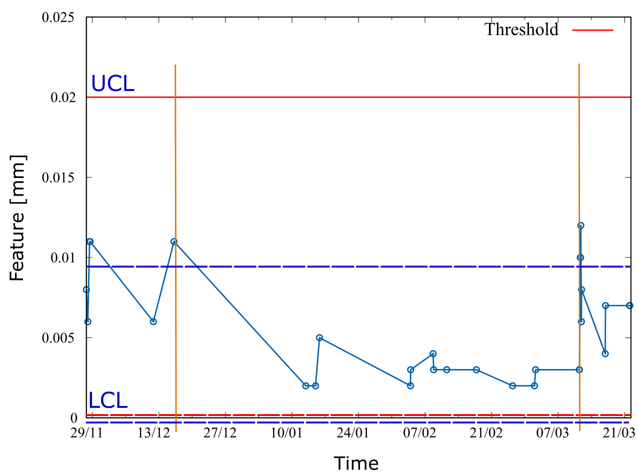

Fig. 4. Detected shifts on a subset of measurements of a sample feature and machine. Each vertical line indicates the point in time where the change is detected. Two moments are identified here, one where there is a tendency towards the target value, which is zero, and one where the measurements start to move away.

3.3.3. SPC rules

Standard SPC rules which include increasing/decreasing patterns and saw tooth patterns alternating below and above the target value, have been used to further characterise the machine behaviour. Amber and red alerts are associated with measurements or sets of measurements that meet the criteria of an SPC rule depending on where in time the rule is triggered. Specifically, the following rules have been implemented:

1. One point outside 3 SDs. Any measurement within 3 SDs from the best machine mean would raise an amber alert, while any measurement outside 3 SDs would raise a red alert.
2. Two out of three points in succession out of 3 SD of the best machine mean. Using a rolling window, any 3 consecutive observations of a set of measurements is tested for this rule. A red alert is raised if the rule is triggered by the any of the 3 most recent measurements, or an amber alert is raised if triggered by any of the other observations in the set. This allows to identify the cases where the rule has been triggered recently or if it has been triggered at all.
3. Four out of five points in succession out of 3SD of the best machine mean. Similar to the rule 2 out of 3, any 4 points in 5 consecutive observations would raise a red alert if triggered on the last three measurements or an amber alert if triggered before.
4. Six points in succession rising or falling. If the last observation is part of an increasing or decreasing trend that is moving away from the nominal, then a red alert will be raised. For any other triggers, an amber alert will be raised.
5. Eight points in succession outside of 1 SD. As in previous rules, any triggers in the last 3 measurements will trigger a red alert, while other measurements will trigger an amber alert.
6. Nine points in succession on the same side. Any trigger of this rule will raise a red alert.
7. Fourteen points in succession alternating above and below the target. As with the previous rule any trigger of this rule will raise a red alert.

All the alerts raised by patterns, shifts and SPC rules are used to generate a summary per machine. Results of all machines are then sorted by severity (red alerts having the highest severity and green the least), allowing to identify the machines that have multiple severe alerts Both the summary and results per rule can be visualised through the visualisation tool for further inspection as it will be shown in the next section.

3.4. Visualisation tool with R

Visualising the statistical results and SPC rules is crucial for manufacturing practitioners to understand the general behaviour of the machines in the manufacturing cell and make decisions based on this [65]. The visualisation dashboard that has been implemented (Fig. 5) has four main visualisation elements which are explained below.

3.4.1. Status panel

As the characterisation of machines for a given feature or set of features is based around the definition of a reference model or best machine, the dashboard provides a quick way for the operator to identify the reference model and related information. This is done through three status boxes at the top of the dashboard (Section A in Fig. 5); a green box, which indicates the best machine, a red box which indicates the worst performing machine and an amber box, which indicates the number of adjustable machines (if any).

3.4.2. Summary statistics tables

To support the results shown in the status panel, the summary statistic tables present the statistical results (mean, SD, variance, max and min) that are computed for each machine and are presented in an ordered manner in section B in Fig. 5), left-hand side table. Details of the adjustable machines are presented as well (right-hand side table).

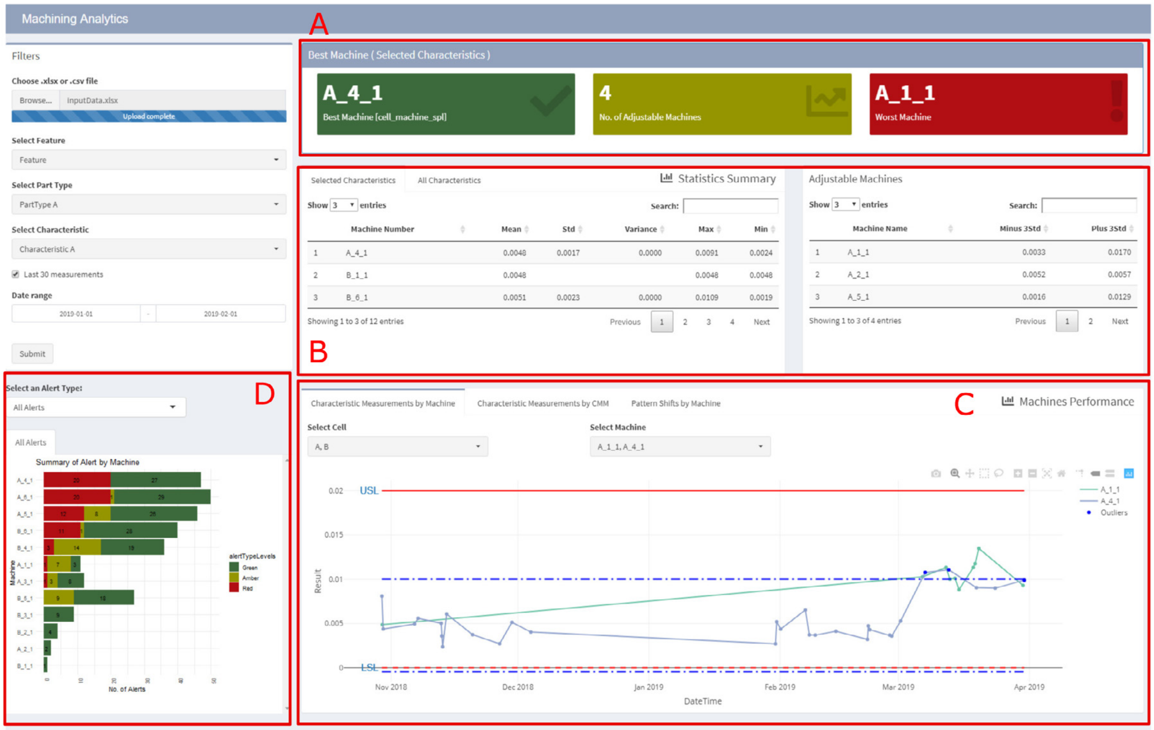

Fig. 5. Visualisation Dashboard consisting of four main visualisations: status panel, statistic tables, alerts summary and control charts.

3.4.3. *Control chart view*

The control chart view (Section C in Fig. 5) allows to visualise the measurements corresponding to a selected feature and time frame for all or sets of machines, highlighting the 3 SD of the best machine (blue dotted line), the outliers (blue dots) and providing the regular elements of a control chart such as the lower and upper control limits (shown in red lines). Data can also be inspected by the CMM machine that performed the measurements, as sometimes there is a consistent error in the measurement machine rather than issues with a particular machine on the manufacturing cell. Finally, it provides a way to visualise change points per machine if the quality practitioner requires further details on the behaviour of a particular machine.

3.4.4. *Alert summary and details view*

The alert summary (Section D in Fig. 5) provides in a first instance a general view of the machines' performance by displaying in order of severity the alerts generated per machine. This allows the operator to identify the machines that are currently more problematic, displaying a large number of red alerts, but also identify which machines could be potentially adjusted as they start showing early signs of a change in the quality of the machined parts. This dashboard feature also allows to filter the summary per rule, which provides further inspection of all the alerts raised by a machine for a particular rule.

4. Industrial case study

The proposed methodology and visualisation tool was tested in a real flexible manufacturing cell at BMW Plant Hams Hall in Birmingham, UK. The UK is the only place where the BMW Group has its three brands BMW, MINI and Rolls-Royce Motor Cars represented by manufacturing operations. The Hams Hall plant in particular manufactures the latest generation of BMW TwinPower Turbo petrol engines, the latest development in the BMW Efficient Dynamics engine family. Every derivative is produced on the basis of a core engine and a modular kit and so there is a high number of parts that are identical across the family. This enables the plant to have a highly standardised and flexible manufacturing network. In this network, core engine components such as crankshafts, cylinder blocks and cylinder heads are machined from raw castings through turning, milling, drilling, polishing and honing processes. The quality of the machined components is constantly monitored throughout the production process and finished parts go to either assembly within the same plant or to other engine manufacturing plants. For this reason, frequent measurement of parts is central to their operation. The flexible cell configuration is similar to the one shown in Fig. 6, where there are between 15 and 20 machines per cell/process, 2 spindles or fixtures per machine and so generally 2 tools per machine. For 8 product types, for example, there can be around 400 measured features, which makes an estimated 96,000 control charts to monitor.

During regular operation of the machines, although appropriate warning limits allow operators to determine the scope of which features need to be reacted to, processes or features which have been affected or shifted but still fall within acceptable limits are often missed. This is directly reflected on the amount of scrap that is produced. Furthermore, the tolerance bands can vary between features and their significance. There are some characteristics which have a range of less than 10 μm and some which are up to 100 μm.

Additional challenges are imposed by the feature measurement variation. Firstly, measurement duration can take up to 90 min for particularly complex processes with 400–500 features on CMMs. This is often why the sampling frequency cannot typically be increased across the whole process. As a result, the lead or 'reference' machine concept is used. The frequency of measure on a single machine within a process group (with multiple machines) is increased and the results of the remaining machines is analysed based on the mean and sparsity of results. Secondly, there is a limited amount of measurement capacity in terms of CMMs because of factors such as the process flow of material through the logistics modules and wash processes. The gauge systems require stringent cleanliness due to the product tolerances (within microns) so they often have to go through high pressure wash systems before measurement (see Fig. 7). This all adds to process time and can reduce capacity. Finally, technologies of the measurement devices also vary, not only with their own resolution and measurement accuracy, but with data accuracy and transfer.

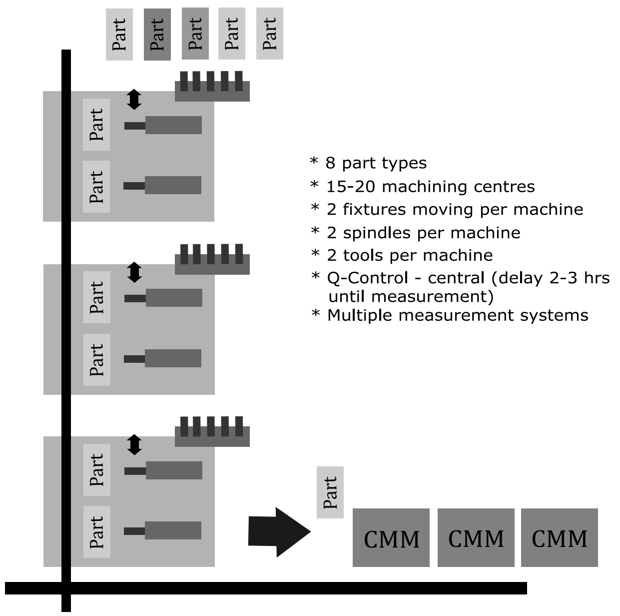

Fig. 6. Typical layout of a Flexible Manufacturing Line at Hams Hall with 15 to 20 machine centres and 2 fixtures and spindles per machine.

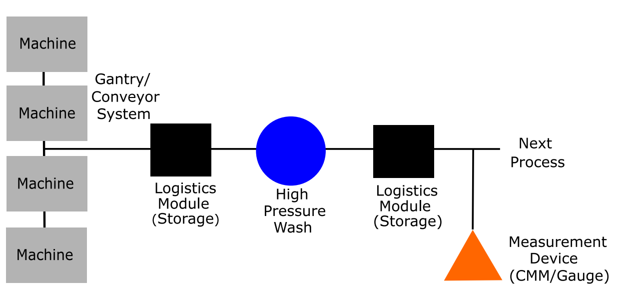

Fig. 7. Selected parts for measurement move through a conveyor system in order to be high-pressure washed before they go through the measurement process.

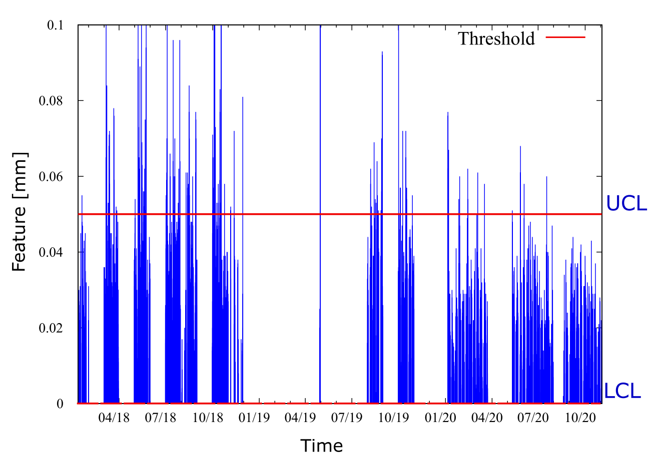

Fig. 8. Measurements of a single feature captured for all monitored machines during the period 2018–2020. For this feature, the upper limit tolerance is indicated with a horizontal red line, and the target value is zero. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To avoid disruption on the production lines, the tool was deployed progressively between January and November 2020. The tool, which was developed using R and shiny libraries, was containerised using Docker to simplify the deployment across the plant. In a first instance,

Table 1
Process capability index (cpk) of three machines before using the analytics tool (all measurements taken in 2018) and after using the tool (period of 11 months in 2020).

Machine number	Before tool	After tool
Machine 1	0.61	2.50
Machine 2	0.43	1.46
Machine 3	0.42	1.54

one installation of the tool was done in January 2020 for the quality specialists to monitor a focused range of part types and support the weekly process capability reviews carried out. The usage of the tool was managed by one of the quality data specialists of the Quality Department at BMW during these weekly reviews, supporting the other specialists to familiarise with its usage. Any characteristics or features which were considered statistically less stable by the tool, were analysed and focused upon during these weekly reviews in addition to existing process control methods. Typically, significant amount of time is spent preparing and performing these reviews to find the right data, make plots and visualisations and interpreting the machine/part type influences before even applying SPC techniques in the evaluations. The tool enabled the team to draw very quick conclusions on where they need to focus their attention. This allowed the focus discussions to be around the technical problems with the machines or processes and encouraging more effective problem resolution activities. After the first month of usage and once the value of reducing this analytic time was recognised, the tool was deployed in a wider set of devices so the team individually began to use the tool more frequently to help steer daily operations for processes which needed more attention. In doing so, operators would select a range of features from their operation, determine the reference machine, and use the “adjustable machines” graphic to help evaluate how they could improve the processes. Once standard practice of using the tool was established, continuous machine adjustments were made to help centre the processes. These adjustments were documented each time and enabled cross shift communication to recognise the best adjustment techniques on the machines. Adjustments included the replacement of worn tools and change in the machine parameters such as offset and thrust. This strengthened the view and approach to using prescriptive analytics for heavily manufacturing processes with large amounts of flexibility and design complexity.

By being able to identify promptly relevant patterns in quality data, processes could be adjusted and centred, keeping part features within the specification limits, improving significantly the performance of the monitored machines. Fig. 8 shows all the measurements taken for a single feature in all machines for a period of two years, from 2018 to 2020. This feature has a unilateral tolerance, where the upper control limit is 0.5. It can be observed in the figure that CMM measurements start to reflect a more stable and controlled process as measurements out of tolerance are less frequent, especially towards the second half of 2020. Inspecting individual machines as shown in Figs. 9 and 10, it can be observed that this tendency is true for all the monitored machines. When comparing the process capability index of machines corresponding to measurements taken in 2018 to those in 2020, it was found that there was a significant improvement across all machines, maintaining a high percentage of produced parts that meet the quality requirements (see Table 1). Initial usage of the tool demonstrated a reduction in 97% of scrap produced for a focused process and group of features within the 11 month period compared the previous year. With these initial results, it can be estimated that 80% of scrap will be reduced from dimensional process faults. Furthermore, with wider use throughout the shop floor, an estimated 500 h of lost process time can be expected to be gained from reduced validation measurements and wasted process steps. This is estimated based on a year’s worth of data from scrap parts with dimensional errors in previous years.

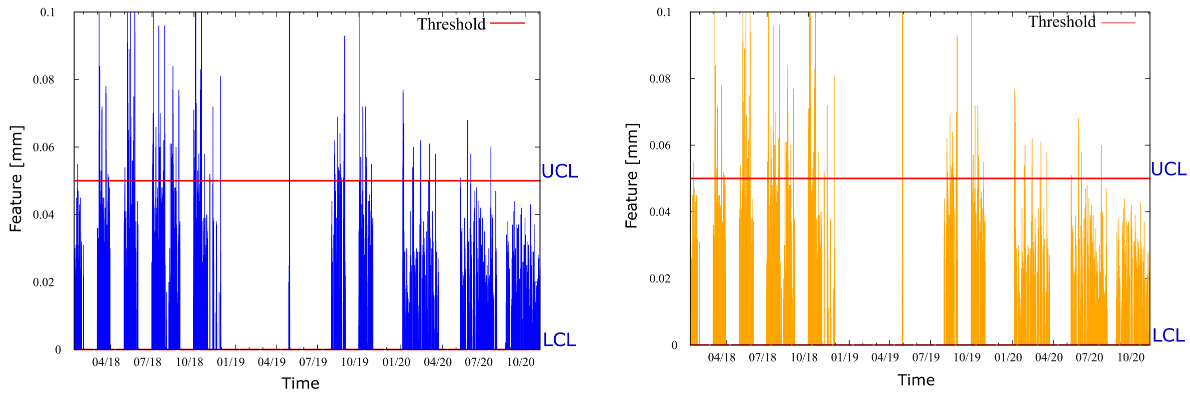

Fig. 9. Measurements of a single feature captured for Machine 1 (left) and Machine 2 (right) during the period 2018–2020. For both machines, the target value is zero and the upper limit tolerance is shown with a red horizontal line.

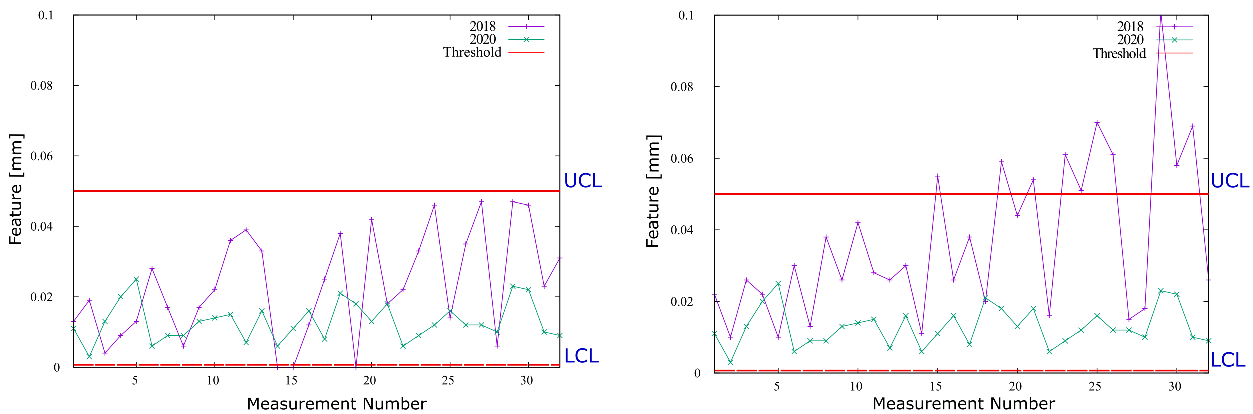

Fig. 10. Thirty two consecutive measurements of a single feature captured for Machine 1 (left) and Machine 2 (right) during the periods 2018 and 2020. As a unilateral tolerance, the target value in both cases is zero.

5. Conclusion and future work

In this work, a methodology for characterising machine performance through part quality data and data analytics is presented. The methodology, which has been integrated into a visualisation tool, allows the quality practitioner to draw out the key geometrical features to analyse and to identify the ones that have adjustable capabilities within their scope. This is crucial, particularly when there is a large number of features to monitor and the level of sensitivity in error is high. By using the tool, the amount of time taken to search the data, inspect control charts and perform the analysis is significantly reduced. This allowed the tool not only to be used for regular practice during weekly quality reviews, but to be used by operators on a daily basis in the shopfloor. The methodology is successful at identifying existing issues as well as at anticipating potential problems that would have been missed otherwise. Through the adjustable machines feature and the documentation of adjustments made by operators, cross shift communication was enabled, setting the basis for a prescriptive approach in a highly flexible manufacturing system. Initial usage of the tool in a flexible manufacturing cell at BMW has demonstrated a reduction of 97% of scrap for the monitored features compared to the previous year. This in the long run will reduce reaction time in following quality control procedures. Furthermore, it will reduce significant scrap costs and ultimately reduce required measurement capacity and enable more output in terms of volume capacity. Based on the scrap produced in 2019 from dimensional process faults, it is estimated that for 2020, 80% of the scrap for the studied set of part types will be reduced, which translates into the reduction of 500 h of process time that would have been expected from validation measurements and wasted process steps. Further work will include the study of a wider use of the tool throughout the shop floor, analysing the usability from the perspective of the shop-floor workers, as well as the incorporation of more advanced machine learning algorithms for pattern recognition. Ultimately prescriptive analytics will be explored to provide an intelligent recommendation of parameter adaptation.

Ethics approval

No ethical approval was needed for this research as no personal data was gathered.

Consent to participate

All participants in the visualisation tool evaluation are employees of BMW Hams Hall Plant and gave informed affirmative consent to the authors to have their evaluation of the tool included in this paper. No personally identifying information was collected.

Consent to publish

This paper presents original work that is not under consideration in any other journal. All authors approved the manuscript and this submission.

Funding

This work was funded by BMW Hams Hall Plant, UK, the ESPRC Impact Acceleration Account, UK grant agreement EP/R511730/1 and the EPSRC Manufacturing on a Shoestring, UK grant agreement EP/R032777/1.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Giovanna Martinez Arellano reports financial support was provided by BMW Hams Hall Plant.

Availability of data and materials

Data cannot be made publicly available due to a confidentiality agreement with BMW Hams Hall Plant. Due to this agreement, the visualisation tool that has been developed as part of this work is not available for public use.

References

[1] A.D. Toni, S. Tonchia, Manufacturing flexibility: A literature review, Int. J. Prod. Res. 36 (6) (1998) 1587–1617, http://dx.doi.org/10.1080/002075498193183.

[2] A. Arbel, A. Seidmann, Performance evaluation of flexible manufacturing systems, IEEE Trans. Syst. Man Cybern. SMC-14 (4) (1984) 606–617, http://dx.doi.org/10.1109/TSMC.1984.6313332.

[3] K. Mahmood, T. Karaulova, T. Otto, E. Shevtshenko, Performance analysis of a flexible manufacturing system (FMS), Procedia CIRP 63 (2017) 424–429, http://dx.doi.org/10.1016/j.procir.2017.03.123, Manufacturing Systems 4.0 – Proceedings of the 50th CIRP Conference on Manufacturing Systems, URL: http://www.sciencedirect.com/science/article/pii/S2212827117302706.

[4] C. Kuo, H. Huang, Failure modeling and process monitoring for flexible manufacturing systems using colored timed Petri nets, IEEE Trans. Robot. Autom. 16 (3) (2000) 301–312, http://dx.doi.org/10.1109/70.850648.

[5] P.A. Marques, C. Cardeira, P. Paranhos, S. Ribeiro, H. Gouveia, Selection of the most suitable statistical process control approach for short production runs: A decision-model, Int. J. Inf. Educ. Technol. (2015) 303–310, http://dx.doi.org/10.7763/IJIET.2015.V5.521.

[6] A.F.B. Costa, P. Castagliola, Effect of measurement error and autocorrelation on the x^{-} chart, J. Appl. Stat. 38 (4) (2011) 661–673, http://dx.doi.org/10.1080/02664760903563627.

[7] C.K. Koh, J.F. Chin, S. Kamaruddin, Modified short-run statistical process cotrol for test and measurement process, Int. J. Adv. Manuf. Technol. 100 (2019) 1531–1548, http://dx.doi.org/10.1007/s00170-018-2776-1.

[8] A. Korzenowski, G. Vidor, G. Vaccaro, C. Ten Caten, Control charts for flexible and multi-variety production systems, Comput. Ind. Eng. 88 (C) (2015) 284–292, http://dx.doi.org/10.1016/j.cie.2015.07.016.

[9] Q.P. He, J. Wang, Statistical process monitoring as a big data analytics tool for smart manufacturing, J. Process Control 67 (2018) 35–43, http://dx.doi.org/10.1016/j.jprocont.2017.06.012, Big Data: Data Science for Process Control and Operations, URL: https://www.sciencedirect.com/science/article/pii/S0959152417301257.

[10] D. Tiuc, G. Draghici, A. Parvu, B. Enache, Consideration about the determination and control of the key characteristics as part of planning quality of the product development process, in: Innovative Manufacturing Engineering 2015, in: Applied Mechanics and Materials, vol. 809, Trans Tech Publications Ltd, 2015, pp. 1269–1274, http://dx.doi.org/10.4028/www.scientific.net/AMM.809-810.1269.

[11] Quality Management Standards, 2015, ISO 9002:2015, International Organisation for Standardisation, 2015.

[12] M. Amini, S. Chang, A review of machine learning approaches for high dimensional process monitoring, in: Proceedings of the 2018 Industrial and Systems Engineering Research Conference, 2018.

[13] C. Herrmann, C. Schmidt, D. Kurle, et al., Sustainability in manufacturing and factories of the future, Int. J. Precis. Eng. Manuf. Green Technol. (2014) 283–292, http://dx.doi.org/10.1007/s40684-014-0034-z.

[14] Y. Kapoor, R. Meyer, B. Meyer, et al., Flexible manufacturing: The future state of drug product development and commercialization in the pharmaceutical industry, J. Pharm. Innov. (2020) http://dx.doi.org/10.1007/s12247-019-09426-z.

[15] R. Cordero, Changing human resources to make flexible manufacturing systems (FMS) successful, J. High Technol. Manag. Res. 8 (2) (1997) 263–275, http://dx.doi.org/10.1016/S1047-8310(97)90006-7, URL: http://www.sciencedirect.com/science/article/pii/S1047831097900067.

[16] D.C. Montgomery, Introduction to Statistical Quality Control, seventh ed., Wiley, 2013.

[17] F.S. Hillier, $\bar{X}$- and R-chart control limits based on a small number of subgroups, J. Qual. Technol. 1 (1) (1969) 17–26, http://dx.doi.org/10.1080/00224065.1969.11980343.

[18] C.P. Quesenberry, SPC Q charts for start-up processes and short or long runs, J. Qual. Technol. 23 (3) (1991) 213–224, http://dx.doi.org/10.1080/00224065.1991.11979327.

[19] P.F. Zantek, Run-length distributions of Q-chart schemes, IIE Trans. 37 (11) (2005) 1037–1045, http://dx.doi.org/10.1080/07408170500232297.

[20] F. He, W. Jiang, L. Shu, Improved self-starting control charts for short runs, Qual. Technol. Quant. Manag. 5 (3) (2008) 289–308, http://dx.doi.org/10.1080/16843703.2008.11673402.

[21] L. Zhang, G. Chen, P. Castagliola, On t and EWMA t charts for monitoring changes in the process mean, Qual. Reliab. Eng. Int. (8) (2009) 933–945, http://dx.doi.org/10.1002/qre.1012.

[22] K. Gu, X. Jia, H. You, S. Zhang, A t-chart for monitoring multi-variety and small batch production run, Qual. Reliab. Eng. Int. 30 (2) (2014) 287–299, http://dx.doi.org/10.1002/qre.1496, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/qre.1496, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qre.1496.

[23] A. Haq, S. Ejaz, M.H. Lee, M. Khan, A new double EWMA-t chart with auxiliary information for the process mean, Qual. Reliab. Eng. Int. 37 (8) (2021) 3381–3394, http://dx.doi.org/10.1002/qre.2923, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/qre.2923, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qre.2923.

[24] W. Jiang, K.-L. Tsui, W.H. Woodall, A new SPC monitoring method: The ARMA chart, Technometrics 42 (4) (2000) 399–410, http://dx.doi.org/10.1080/00401706.2000.10485713, URL: https://www.tandfonline.com/doi/abs/10.1080/00401706.2000.10485713.

[25] Z. Wu, M. Shamsuzzaman, E.S. Pan, Optimization design of control charts based on taguchi's loss function and random process shifts, Int. J. Prod. Res. 42 (2) (2004) 379–390, http://dx.doi.org/10.1081/00207540310001614169.

[26] Z. Wu, J. Jiao, M. Yang, Y. Liu, Z. Wang, An enhanced adaptive CUSUM control chart, IIE Trans. 41 (7) (2009) 642–653, http://dx.doi.org/10.1080/07408170802712582.

[27] A. Amiri, S. Allahyari, Change point estimation methods for control chart postsignal diagnostics: A literature review, Qual. Reliab. Eng. Int. 28 (7) (2012) 673–685, http://dx.doi.org/10.1002/qre.1266, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/qre.1266, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qre.1266.

[28] D.M. Hawkins, P. Qiu, C.W. Kang, The changepoint model for statistical process control, J. Qual. Technol. 35 (4) (2003) 355–366, http://dx.doi.org/10.1080/00224065.2003.11980233.

[29] R. Cummings, S. Krehbiel, Y. Mei, R. Tuo, W. Zhang, Differentially private change-point detection, in: Advances in Neural Information Processing Systems 31 NeurIPS, 2018, pp. 10825–10834.

[30] H. Jin, G. Yin, B. Yuan, F. Jiang, BayesIan hierarchical model for change point detection in multivariate sequences, Technometrics (2021) 1–10, http://dx.doi.org/10.1080/00401706.2021.1927848.

[31] Y. Ding, L. Zeng, S. Zhou, Phase I analysis for monitoring nonlinear profiles in manufacturing processes, J. Qual. Technol. 38 (3) (2006) 199–216, http://dx.doi.org/10.1080/00224065.2006.11918610.

[32] S. Yasui, H. Noguchi, Y. Ojima, A Robust Detection Procedure for Multiple Change Points of Linear Trends, 2012, pp. 197–206, http://dx.doi.org/10.1007/978-3-7908-2846-7-14, 2010 10th International Workshop on Intelligent Statistical Quality Control ; Conference date: 18-08-2010 Through 20-08-2010.

[33] D.M. Hawkins, Q. Deng, A nonparametric change-point control chart, J. Qual. Technol. 42 (2) (2010) 165–173, http://dx.doi.org/10.1080/00224065.2010.11917814.

[34] W. Li, C. Zhang, F. Tsung, Y. Mei, Nonparametric monitoring of multivariate data via KNN learning, Int. J. Prod. Res. 59 (20) (2021) 6311–6326, http://dx.doi.org/10.1080/00207543.2020.1812750.

[35] L.A. Jones-Farmer, W.H. Woodall, S.H. Steiner, C.W. Champ, An overview of phase I analysis for process improvement and monitoring, J. Qual. Technol. 46 (3) (2014) 265–280, http://dx.doi.org/10.1080/00224065.2014.11917969.

[36] T. Zan, Z. Liu, Z. Su, M. Wang, X. Gao, D. Chen, Statistical process control with intelligence based on the deep learning model, Appl. Sci. 10 (308) (2020) http://dx.doi.org/10.3390/app10010308.

[37] J. Wang, A. Kochhar, R. Hannam, Pattern recognition for statistical process control charts, Int. J. Adv. Manuf. Tehcnol. 14 (1998) 99–109, http://dx.doi.org/10.1007/BF01322218.

[38] M.B.C. Khoo, Design of runs rules schemes, Qual. Eng. 16 (1) (2003) 27–43, http://dx.doi.org/10.1081/QEN-120020769.

[39] T. Zan, Z. Liu, H. Wang, M. Wang, X. Gao, Control chart pattern recognition using the convolutional neural network, J. Intell. Manuf. 31 (2020) 703–716, http://dx.doi.org/10.1007/s10845-019-01473-0.

[40] E. García, R. Peñabaena-Niebles, M. Jubiz-Diaz, A. Perez-Tafur, Concurrent control chart pattern recognition: A systematic review, Mathematics 10 (6) (2022) http://dx.doi.org/10.3390/math10060934.

[41] R.-S. Guh, Integrating artificial intelligence into on-line statistical process control, Qual. Reliab. Eng. Int. 19 (1) (2003) 1–20, http://dx.doi.org/10.1002/qre.510, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/qre.510, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qre.510.

[42] J. Yu, L. Xi, A neural network ensemble-based model for on-line monitoring and diagnosis of out-of-control signals in multivariate manufacturing processes, Expert Syst. Appl. 36 (1) (2009) 909–921, http://dx.doi.org/10.1016/j.eswa.2007.10.003, URL: http://www.sciencedirect.com/science/article/pii/S0957417407004939.

[43] S. He, Z. He, G. Wang, Online monitoring and fault identification of mean shifts in bivariate processes using decision tree learning techniques, J. Intell. Manuf. 24 (2013) 25–34, http://dx.doi.org/10.1007/s10845-011-0533-5.

[44] M.S. Reis, G. Gins, Industrial process monitoring in the big data/industry 4.0 era: from detection, to diagnosis, to prognosis, 2017.

[45] A. Diez-Olivan, J.D. Ser, D. Galar, B. Sierra, Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0, Inf. Fusion 50 (2019) 92–111, http://dx.doi.org/10.1016/j.inffus.2018.10.005, URL: http://www.sciencedirect.com/science/article/pii/S1566253518304706.

[46] Y. Liang, W. Li, X. Lu, S. Wang, Fog computing and convolutional neural network enabled prognosis for machining process optimization, J. Manuf. Syst. 52 (2019) 32–42, http://dx.doi.org/10.1016/j.jmsy.2019.05.003, URL: https://www.sciencedirect.com/science/article/pii/S0278612519300226.

[47] A. Azadeh, J. Zeynali, Design and development of an integrated quality control system, Int. J. Prod. Qual. Manag. 3 (2) (2008) 183–205, http://dx.doi.org/10.1504/IJPQM.2008.016564.

[48] C.-T. Huang, T.-M. Yeh, W.-T. Lin, B.-T. Lee, A fuzzy AHP-based performance evaluation model for implementing SPC in the Taiwanese LCD industry, Int. J. Prod. Res. 47 (18) (2009) 5163–5183, http://dx.doi.org/10.1080/00207540801935608.

[49] L. Guerra, S.D. Sousa, E.P. Nunes, Statistical process control automation in the final inspection process: An industrial case study, in: 2016 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM, 2016, pp. 876–880, http://dx.doi.org/10.1109/IEEM.2016.7798002.

[50] J. Schmitt, J. Bönig, T. Borggräfe, G. Beitinger, J. Deuse, Predictive model-based quality inspection using machine learning and edge cloud computing, Adv. Eng. Inf. 45 (2020) 101101, http://dx.doi.org/10.1016/j.aei.2020.101101, URL: https://www.sciencedirect.com/science/article/pii/S1474034620300707.

[51] R.S. Peres, X. Jia, J. Lee, K. Sun, A.W. Colombo, J. Barata, Industrial artificial intelligence in industry 4.0 - systematic review, challenges and outlook, IEEE Access 8 (2020) 220121–220139, http://dx.doi.org/10.1109/ACCESS.2020.3042874.

[52] L. Rychener, F. Montet, J. Hennebert, Architecture proposal for machine learning based industrial process monitoring, Procedia Comput. Sci. 170 (2020) 648–655, http://dx.doi.org/10.1016/j.procs.2020.03.137, The 11th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops, URL: http://www.sciencedirect.com/science/article/pii/S1877050920305925.

[53] J. Moyne, J. Iskandar, Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing, Processes 5 (3) (2017) 39, http://dx.doi.org/10.3390/pr5030039.

[54] L. Stojanovic, M. Dinic, N. Stojanovic, A. Stojadinovic, Big-data-driven anomaly detection in industry (4.0): An approach and a case study, in: 2016 IEEE International Conference on Big Data, Big Data, 2016, pp. 1647–1652, http://dx.doi.org/10.1109/BigData.2016.7840777.

[55] Y. Cohen, M. Faccio, F. Pilati, X. Yao, Design and management of digital manufacturing and assembly systems in the industry 4.0 era, Int. J. Adv. Manuf. Technol. 105 (2019) 3565–3577, http://dx.doi.org/10.1007/s00170-019-04595-0.

[56] Y. Cohen, G. Singer, A smart process controller framework for industry 4.0 settings, J. Intell. Manuf. (2021) http://dx.doi.org/10.1007/s10845-021-01748-5.

[57] Y. Liao, F. Deschamps, E. de Freitas Rocha Loures, L.F. Pierin Ramos, Past, present and future of industry 4.0 - a systematic literature review and research agenda proposal, Int. J. Prod. Res. 55 (12) (2017) 3609–3629, http://dx.doi.org/10.1080/00207543.2017.1308576.

[58] M. Quiñones-Grueiro, A. Prieto-Moreno, C. Verde, O. Llanes-Santiago, Data-driven monitoring of multimode continuous processes: A review, Chemometr. Intell. Lab. Syst. 189 (2019) 56–71, http://dx.doi.org/10.1016/j.chemolab.2019.03.012, URL: https://www.sciencedirect.com/science/article/pii/S0169743918306774.

[59] C. Leys, C. Ley, O. Klein, P. Bernard, L. Licata, Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median, J. Exp. Soc. Psychol. 49 (4) (2013) 764–766, http://dx.doi.org/10.1016/j.jesp.2013.03.013, URL: http://www.sciencedirect.com/science/article/pii/S0022103113000668.

[60] J. Miller, Short report: Reaction time analysis with outlier exclusion: Bias varies with sample size, Q. J. Exper. Psychol. Sect. A 43 (4) (1991) 907–912, http://dx.doi.org/10.1080/14640749108400962.

[61] F. Kadri, F. Harrou, S. Chaabane, Y. Sun, C. Tahon, Seasonal ARMA-based SPC charts for anomaly detection: Application to emergency department systems, Neurocomputing 173 (2016) 2102–2114, http://dx.doi.org/10.1016/j.neucom.2015.10.009, URL: https://www.sciencedirect.com/science/article/pii/S0925231215014654.

[62] D.R. Cox, A. Stuart, Some quick sign tests for trend in location aNd dispersion, Biometrika 42 (1–2) (1955) 80–95, http://dx.doi.org/10.1093/biomet/42.1-2.80.

[63] A.F. Militino, M. Moradi, M.D. Ugarte, On the performances of trend and change-point detection methods for remote sensing data, Remote Sens. 12 (6) (2020) http://dx.doi.org/10.3390/rs12061008, URL: https://www.mdpi.com/2072-4292/12/6/1008.

[64] R. Killick, P. Fearnhead, I.A. Eckley, Optimal detection of changepoints with a linear computational cost, J. Amer. Statist. Assoc. 107 (500) (2012) 1590–1598, http://dx.doi.org/10.1080/01621459.2012.737745.

[65] M.O.A. Abu-Shawiesh, N. Saeed, Monitoring process variability using decile mean standard deviation, Qual. Reliab. Eng. Int. 38 (1) (2022) 501–515, http://dx.doi.org/10.1002/qre.2996, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/qre.2996, URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qre.2996.

Acronyms

ANN: Artificial Neural Network
CMM: Coordinate Measuring Machine
cpk: Process capability index
CUSUM: Cumulative Sum
EWMA: Exponentially Weighted Moving Average
FMS: Flexible Manufacturing Systems
IIND: Independent and Identically Normally Distributed
IT: Information Technology
JIT: Just-in-time
MAD: Median Absolute Deviation
ML: Machine Learning
PELT: Pruned Exact Linear Time
SD: Standard Deviation
SPC: Statistical Process Control