



HAL
open science

Application of Machine Learning in Statistical Process Control Charts: a survey and perspective

Phuong Hanh Tran, Adel Ahmadi Nadi, Thi Hien Nguyen, Kim Duc Tran,
Kim Phuc Tran

► **To cite this version:**

Phuong Hanh Tran, Adel Ahmadi Nadi, Thi Hien Nguyen, Kim Duc Tran, Kim Phuc Tran. Application of Machine Learning in Statistical Process Control Charts: a survey and perspective. Control Charts and Machine Learning for Anomaly Detection in Manufacturing., 2021, Springer Series in Reliability Engineering., 10.1007/978-3-030-83819-5_2 . hal-03544308

HAL Id: hal-03544308

<https://hal.science/hal-03544308v1>

Submitted on 26 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Application of Machine Learning in Statistical Process Control Charts: a survey and perspective

Phuong Hanh Tran¹, Adel Ahmadi Nadi^{2,4}, Thi Hien Nguyen^{1,3}, Kim Duc Tran¹, and Kim Phuc Tran^{*4}

²Department of Statistics, Ferdowsi University of Mashhad, P. O. Box 1159, Mashhad 91775, Iran

³Laboratoire AGM, UMR CNRS 8088, CY Cergy Paris Université, 95000 Cergy, France. Email: thi-hien.nguyen5@cyu.fr

⁴University of Lille, ENSAIT, GEMTEX, F-59000 Lille, France,

*Corresponding author. Email: kim-phuc.tran@ensait.fr

June 15, 2021

Abstract

Over the past decades, control charts, one of the essential tools in Statistical Process Control (SPC), have been widely implemented in manufacturing industries as an effective approach for Anomaly Detection (AD). Thanks to the development of technologies like the Internet of Things and Artificial Intelligence (AI), Smart Manufacturing (SM) has become an important concept for expressing the end goal of digitization in manufacturing. However, SM requires a more automatic procedure with capabilities to deal with huge data from the continuous and simultaneous process. Hence, traditional control charts of SPC now find difficulties in reality activities including designing, pattern recognition, and interpreting stages. Machine Learning (ML) algorithms have emerged as powerful analytic tools and great assistance that can be integrating to control charts of SPC to solve these issues. Therefore, the purpose of this chapter is first to presents a survey on the applications of ML techniques in the stages of designing, pattern recognition, and interpreting of control charts respectively in SPC especially in the context of SM for AD. Second, difficulties and challenges in these areas are discussed. Third, perspectives of ML techniques-based control charts for AD in SM are proposed. Finally, a case study of an ML-based control chart for bearing failure AD is also provided in this chapter.

1 Introduction

Together with the blooming flourish rapidly of data, including velocity, volume, and variety, anomaly detection (AD) has become a hot topic in recent years. The important role of AD has demonstrated throughout various studies in numerous different disciplines such as emergency hospital systems (Kadri et al.¹), traffic measurement (Münz and Carle²), credit card fraud detection (Tran et al.³), manufacturing industry (Tran et al.⁴; Tran and Heuchenne⁵). According to Chandola et al.⁶, AD has seen as a term concern to find the instances that do not well conform to a defined notion of normal behavior. These instances are called anomalies or outliers or interchangeably. The beginning of the 19th century is considered as the milestones of the AD issue that has been dealt with by the statistical science community (Edgeworth⁷). The requirement for early detection of anomalies in the process is necessary to ensure system performance and save time as well as cost for an organization.

It worth mentioning that statistical process control (SPC) is an essential approach for AD that is widely applied in industry. The aim of this approach is to monitor and reduce variation in the process as soon as possible to guarantee high product quality at a minimal cost. In particular, the control chart, one of the fundamental tools of SPC first introduced by Shewhart⁸ has been an effective tool to detect changes and anomalies of characteristics in the procedure. The contribution of the control chart is based on the idea to gives the producers a simple graphical tool for controlling production, i.e. having correction activities in a timely manner. This allows them to keep production centered on its target and to maintain its dispersion within the specified tolerance interval. However, numerous studies show that the implementation of control charts meets some disadvantages in particular situations including designing (Alwan⁹; Noorossana and Vaghefi¹⁰; Costa and Castagliola¹¹; Leoni et al.¹²; Vanhatalo and Kulahci¹³), trend recognition (Guh and Hsieh¹⁴; Swift and Mize¹⁵; Guo and Dooley¹⁶; Miao and Yang¹⁷; Zan et al.¹⁸), and interpreting (Wang and Chen¹⁹; Low et al.²⁰; S. T. A. Niaki and Abbasi²¹) of control chart. A more specific discussion is presented as follows.

It is important to note that a disadvantage of traditional control charts have been discussed in the designing stage. One of the principles in designing a control chart by statistical traditional methods is that it has to under an assumption in which samples of the observed process are normally, independently, and identically distributed (i.i.d. assumption). For example: in the case of univariate process, this implies that the observed in-control process has a steady-state and is characterized by two fixed parameters as mean μ and standard deviation σ . They also lie on an assump-

tion that the main parameters are known or estimated from the historical data. However, this approach faces difficulties in some real activities situations of industry process when considering in the new context as dynamic behavior environment or sampling regularly. Firstly, the normal population distribution assumption is unreal in many cases. Secondly, a variety of researches (Alwan⁹; Noorossana and Vaghefi¹⁰; Costa and Castagliola¹¹; Leoni et al.¹²; Vanhatalo and Kulahci¹³) showed the developed control charts using the assumption of independent observations have been enormously influenced by the presence of autocorrelation. Finally, the complex industry procedure could be dominated by various variables and it is impossible to know the covariance relationships before. This leads to false alarms appear many time. Therefore, efforts to develop advanced control chart using ML in the mentioned cases are necessary.

Besides, control chart pattern recognition (CCPR) is an important problem in SPC. A control chart is used for detecting whether a process is in control or out of control. But one out-of-control state is found and is eliminated, it is necessary to have an observation, i.e., abnormal pattern recognition to well monitor the behavior of the process in the future. Numerous studies focus on CCPR issue from the middle of 1980s (Western²²; Swift²³). The aim of the CCPR task is to diagnose nine common abnormal patterns, i.e. unnatural patterns in the process including upward trend, downward trend, upward shift, downward shift, cycles, runs, stratification, freak patterns, and freak points (Shewhart²⁴). This activity in order to find out and prevent potential causes as soon as possible. CCPR can be performed by quality engineers in small production systems. However, along with the development of manufacturing systems especially SM, sensors are deployed everywhere with huge data sources to be collected and monitored, the application of Machine Learning (ML) to automating this task is an irreversible trend. Miao and Yang¹⁷ reveal that the analysis of the statistical characteristics and shape features of the control chart pattern contribute to recognizing unreal patterns of the process through the relevant algorithm was classified. However, the application of DL methods to automatically extract features from the control chart has proven superior in the ability to recognize patterns, see Zan et al.¹⁸ for more details. Since then, efforts in applying Deep Learning (DL) in this field are a very important research direction.

Finally, a very important issue that needs attention in SPC is the interpretation of out-of-control signals. Traditional univariate control charts have played a significant role in the literature to monitor the characteristic processes for ensuring the quality of the system. However, in real activities of industry, the truth is that the process was dominated by various characteristics in some cases. This issue was often solved by the way of using different univariate control charts. But this would lead to false alarms

when these characteristics have a high correlation or sampling in a short duration. Therefore, it is necessary to collect and monitor multivariate variables simultaneously, i.e. using multivariate statistical process control (MSPC). Hotelling's T^2 chart (Hotelling²⁵), Multivariate Cumulative Sum (MCUSUM) chart (Woodall and Ncube²⁶), and Multivariate Exponentially Weighted Moving Average (MEWMA) chart (Lowry et al.²⁷) are common multivariate control charts of MSPC used to solve the quality control problems. However, a challenge of these traditional multivariate control charts is that they are just only able to detect a shift in the process mean vector, i.e., out-of-control signals of the process. It is impossible to indicate which variable(s) or a group of variables is responsible for out-of-control signals of the process. Moreover, the MSPC requires more rapid identification in comparison with a univariate process that is beyond the capacity of traditional multivariate control charts. The interpretation of out-of-control signals can be considered a classification problem in ML. Therefore, the application of ML to develop methods to automatically interpret the out-of-control signals in the multivariate control charts has attracted a lot of efforts from researchers (Diren et al.²⁸).

In short, thanks to the appearance of ML methods, these difficulties are solved. The application of ML in control charts is a new approach that is overcome these previous disadvantages or issues. Swift²³ and Shewhart²⁴ have seen as the pioneer researchers published ideas combining ML in a control chart. Recently, many pieces of research showed that recognition control-based new ML algorithms have performance better than one based traditional statistical methods as well as conduct to estimate pattern parameters (Guh and Hsieh¹⁴; Guh and Tannock²⁹; Wu and Yang³⁰). Besides, numerous authors also showed that ML methods are useful techniques applied to control charts to tackle the issues in the interpreting stage (Wang and Chen¹⁹; Low et al.²⁰; S. T. A. Niaki and Abbasi²¹; Cheng and Lee³¹). Due to the various advantages of integrating ML techniques to control charts in SPC, we would like to encourage more studies to consider this approach. This can be seen as the alternative one to overcome the above limitations of traditional control charts. However, this is a lack of researches that focuses to give a general picture of these issues in literature. Therefore, the main objective of our chapter is to fill this gap. The remainder of this chapter is organized as follows. Section 2 briefly reviews the design of control chart-based ML methods. Section 3 makes a literature review relevant to CCPR. Section 5 presents the recent studies about the issue of the interpreting-based ML of control charts. Difficulties and challenges in these areas are discussed in Section 5. Section 6 proposed perspectives for ML techniques-based control charts for AD in SM. An experiment for a case study is proposed in Section 7. Finally, concluding remarks of the study are outlined in Section 8.

2 Machine Learning (ML) techniques based Control Charts for process monitoring

Control charts have been developed and applied a lot (see Fig 6, taken from Web of Science), major publications in the fields of the engineering industry. Control charts provide a simple method that can be used to indicate whether the process is stable or not (in control or out-of-control). In detail, it is a chronological graph whose dots represent the tracking of a characteristic of the process. A horizontal line represents the central value (the average). The lower control limit (LCL) and the upper control limit (UCL) are represented by two horizontal lines on either side of the mean. The values of a measured characteristic must be within these limits; otherwise, the process is out of control and must be examined. The main benefits of control charts are: 1) they increase productivity by the proportion of "good product" and decrease costs because there is less waste; 2) they give an estimate of the central tendency of the characteristic of interest, its variability, and the limits within which it varies; 3) control charts assist in the evaluation of the performance of a measurement system. One of the major advantages of the control card is its ease of construction and use, an operator or engineer familiar with the technique of control charts can, in general, diagnose the cause of a problem. However, in order for the control chart to be a reliable and effective indicator of the status of the process, the production using the control chart should pay special attention to the type of chart used.

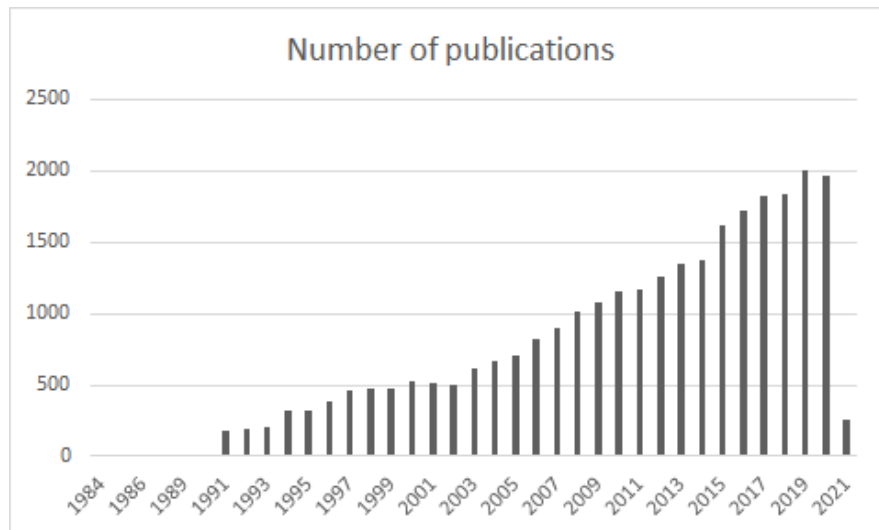


Figure 1: Number of publications on Control Charts from 1984

ML is a domain of Artificial Intelligence (AI), which consists of pro-

programming algorithms to learn automatically from data and experiences or by interaction with the environment. What makes ML really useful is the fact that the algorithm can "learn" and adapt its results based on new data without any a priori programming. There are three main branches: supervised learning, unsupervised learning, and reinforcement learning. The algorithm of supervised learning is to find correlations between input data (explanatory variables) and output data (predictable variables), for then infer the knowledge extracted on inputs with unknown outputs. Different from supervised learning, the technique of unsupervised learning must discover by itself the structure according to the data, which has only one dataset collected as input. This technique is used to divide data into groups of homogeneous items/datapoint. Finally, Reinforcement learning is an area of machine learning concerned with how to make a sequence of decisions. In literature and practice, many researchers have combined techniques of ML and control charts. As mentioned above, by the ease of use of controls charts and the wide application of ML, this combination is increasingly researched and applied. This is because many types of problems that are arising during the implementation of control charts in nowadays complex processes can be effectively solved with the help of ML approaches (see for example Kang et al.³² and Qiu and Xie³³). One of the main contributions of applying ML techniques in designing control charts is that the modern (production, insurance, healthcare, and etc) processes generate huge data sets with a large degree of diversity by means of modern measurement systems like sensors. In such situations, the traditional statistical monitoring methods fail to handle the monitoring procedure of such processes while ML techniques are able to provide impressive results (Weese et al.³⁴). This section will summarize the most common techniques for designing control charts with ML methods for process monitoring.

2.1 Kernel-based Learning Methods

Kernel-based learning methods such as the Support Vector Machine (SVM) algorithm are extensively used and play major roles in the SPC activities, both in developing control charts and recognition of abnormal patterns, due to their remarkable solutions for existing problems. In brief, kernels have been applying in the ML area because, when it is difficult to do a task in the original problem space, the kernel method enables the practitioner to transform the problem space into another in which they can work easier. Recently, Apsemidis et al.³⁵ provided a comprehensive review on about 90 articles after 2002 that include the combination of kernel-based approaches with other ML techniques. Mashuri et al.³⁶ proposed a $Tr(R2)$ control chart based on the squared correlation matrix with the trace operator and used the kernel density estimation method to calculate the better control limit for the proposed chart. Chinnam³⁷ demonstrates that SVMs can be extremely

effective in minimizing both Type-I errors and Type-II errors and in detecting shifts in both the non-correlated processes or autocorrelated processes. A comparison of SVM and Neural Network (NN) for drug/nondrug classification has been done by Byvatov et al.³⁸ and it was demonstrated that the SVMs classifier yielded slightly higher prediction accuracy than NN. By the efficiency of SVMs, many researchers used this technique based on control charts. For example, Li and Jia³⁹ proposed a SVMs based model for fault identification in MEWMA control charts, they examined the effects of SVM parameters on classification performance and provide a SVM parameter optimization method.

Although the kernel-based ML algorithms are mainly applied as classifiers for dividing data into two or more classes, in most of SPC implementations training data from one class (normal state) are only available and there is no information about the other class (abnormal state). This situation may arise from several reasons such as the general difficulties (lack of resources or time or cost) or even impossibility of collecting enough observations for the abnormal class to learn the ML algorithm (Camci and Chinnam⁴⁰). To handle such situations, one-class classifiers are introduced. One-class classifier just learns from the normal training data and labelled the newly encountered data as in-class or out-of-class observations. Several one-class classifiers have been developed by researchers, while support vector data description (SVDD), the k nearest neighbor data description (KNNDD), and K means data description (KMDD) one-class classifiers were only used to develop control charts. One of the first studies in this domain was conducted by Sun and Tsung⁴¹ who designed a kernel distance-based chart (K -chart) using SVDD algorithm, as a modified version of the original SVM for solving one-class classification problems, and concluded that the K chart outperforms conventional charts when the data distribution departs from normality. This work improved by Ning and Tsung⁴² for non-normal process data. Sukchotrat et al.⁴³ developed a K chart that integrates a traditional control chart technique with a KNNDD algorithm, one of the one-class classification algorithms. Later, to examine the feasibility of using one-class classification-based control charts to handle multivariate and autocorrelated processes, Kim et al.⁴⁴ developed a K chart that uses original observations instead of residuals to monitor autocorrelated multivariate processes. Throughout a simulation study, they showed that the K charts outperformed the T^2 control charts, and the performance K charts is not significantly affected by the degrees of autocorrelation. Gani and Limam⁴⁵ examined the performance of the K chart and KNNDD chart through a real industrial application. They investigated the effectiveness of both charts in detecting out-of-control observations using the average run length (ARL) criterion. The results of this study show that the K chart is sensitive to small shifts in the mean vector, whereas the KNNDD chart is sensitive to moderate shifts in the mean vector. In addition, Gani and Limam⁴⁶ introduced a

new chart using the KMDD algorithm and reported that their chart has a better performance in detecting small shifts of mean vector based on ARL than the K chart and KNNDD chart. To improve the performance of K -charts, Maboudou-Tchao et al.⁴⁷ used a one-class SVM technique based on the SVDD method for monitoring the mean vector based on Mahalanobis kernel. They used the Mahalanobis kernel as an alternative for Gaussian kernel and showed that the proposed method is more sensitive than SVDD using Gaussian kernel for detecting shifts in the mean vectors of three different multivariate distributions. They also demonstrated that the proposed method outperforms Hotelling's T^2 chart in multivariate normal cases.

Zhang et al.⁴⁸ developed a general monitoring framework for detecting location shifts in complex processes using the SVM model and multivariate EWMA chart. Later, Wang et al.⁴⁹ developed SVM-based one-sided control charts to monitor a process with multivariate quality characteristics. They used the differential evolution (DE) algorithm to obtain the optimal parameters of the SVM model by minimizing mean absolute error. In this study, the performance of the control charts is investigated using a multivariate normal distribution and two non-normal distributions by considering different process shift scenarios. In addition, through an ARL analysis using the Monte Carlo simulations, they showed that the proposed chart has better performance than the distance-based control charts based on SVM studied by He et al.⁵⁰. Recently, Maboudou-Tchao⁵¹ introduced a least-squares one-class SVM (LS-OCSVM) for monitoring the mean vector of processes. They counted several advantages of their proposed monitoring approach over the existing SVDD chart provided by Sun and Tsung⁴¹ and Hotelling's T^2 chart in terms of simplicity in computation and design, flexibility in implementation, and superiority in performance. For example, they claimed that the LS-OCSVM method can be easily extended to online monitoring. This feature is very beneficial when we are facing a large-scale training dataset that updates over time. The SVDD method uses a batch learning phase in which we learn on the entire training set and generate the best model at once. If new additional training data arrive, SVDD must be retrained from scratch. Using SVM techniques based on control charts to have a better performance can be found at many works, see for example, He et al.⁵⁰, Salehi et al.⁵², Hu and Zhao⁵³, Gani et al.⁵⁴, Sukchotrat et al.⁵⁵, Kakde et al.⁵⁶, Jang et al.⁵⁷.

Regression analysis is a technique of supervised ML. It is based on the basic principles of physics that help predict the future from current data. It also helps to find the correlation between two variables to define the cause and effect relationship. However, there are different forms of regression, ranging from linear regression and complex regression. One of the regression variants which yields very good results is the support vector regression (SVR) method. This technique has been applied a lot in the construction

of control charts, especially when the process variables are highly auto-correlated. For example, Issam and Mohamed⁵⁸ apply the SVR method for the construction of a residuals Multivariate Cumulative Sum (MCUSUM) control chart to monitoring changes in the process mean vector. This chart is designed to detect small shifts in the process parameters and it performed better than the time series based control chart because it can handle non-linear relation between the controlled variables and do not use any restrictive assumption. In 2013, Du et al.⁵⁹ proposed one new Minimal Euclidean Distance (MED) based control chart for recognizing the mean shifts of auto-correlated processes. They also used SVR to predict the values of a variable in time series. The numerical results showed that the MED chart outperformed those of some statistics-based charts and the neural-networks-based (NN) control scheme for the small process mean shifts. Another example of a combination of SVR technique and Control charts, Gani et al.⁵⁴ designed a SVR-chart which using SVR to construct robust control charts for residuals. By comparing the behavior of Average Run Length (ARL), the authors showed that the efficiency of this chart is better than ordinary least squares (OLS), and the partial least squares method.

Besides the above-mentioned supervised learning methods, unsupervised learning algorithms are another type of ML algorithms that applied to analyze and cluster unlabelled datasets. Clustering is one of the most important unsupervised ML techniques, in which similar traits are used to make a prediction. The algorithm measures the proximity between each element based on defined criteria. K-Means is the most popular method of grouping input data, which allows you to set the value of K and order the data according to that value. The aim of the study of Silva et al.⁶⁰ is to apply the u-chart to find out the number of clusters in the K-means method on Automatic Clustering Differential Evolution (ACDE) in order to identify the behavior patterns and relations between the different attributes. These results in this work showed that the use of an u-chart increases the performance of ACDE. Another example of application clustering technique based on control charts in medicine, Thirumalai et al.⁶¹ gave a prediction of diabetes disease for people of various age groups and genders by using cost optimization and control chart.

2.2 Dimensionality Reduction

For a given data, the higher the number of variables, the more complex the results will be, which will make it difficult to consolidate the data. Dimensionality reduction is considered a method of ML to overcome this difficulty. Instead of studying the data involved in a grand dimension, the technique of dimensionality reduction is to replace it with data in a smaller dimension. Roughly speaking, principal components analysis (PCA) is one of

the most important methods of dimensionality reduction that transforms a large dataset of (possibly) correlated observations into a smaller data set of uncorrelated observations by minimizing information loss. Developing control charts based on the PCA method has been widely investigated in the literature. For example, Stefatos and Hamza⁶² introduced a robust multivariate statistical control chart using the Kernel PCA (KPCA) method. They reported that the new chart is robust to outliers detection and performs better than some existing multivariate monitoring and control charts. Phaladiganon et al.⁶³ presented non-parametric PCA technique, kernel density estimation, and bootstrapping to establish the control limits of control charts that. The proposed non-parametric PCA control charts performed better than the parametric PCA control charts in non-normal situations through the behavior of average run length. The PCA's technique is also used in Kullaa⁶⁴, the author showed that the sensitivity of the control chart to damage was substantially increased by further dimensionality reduction applying the principal component analysis. Applying this technique, Lee et al.⁶⁵ developed a new KPCA-based non-linear process monitoring technique for tackling the nonlinear problem. Base on T^2 and squared prediction error (SPE) charts in the feature space, KPCA was applied to fault detection in two example systems: a simple multivariate process and the simulation benchmark of the biological waste-water treatment process. These examples demonstrated that the proposed approach can effectively capture nonlinear relationships in process variables and that, when used for process monitoring, it shows better performance than linear PCA. Using Hotelling's T^2 statistic, Ahsan et al.⁶⁶ implemented the KPCA method for simultaneously monitoring mixed (continuous and categorical) quality characteristics. In this study, it is demonstrated that the KPCA-based control charts have a great performance in terms of successful detection of the out-of-control observations in comparison with the conventional PCA mix charts discussed in Ahsan et al.⁶⁷. Another study in the area of monitoring procedures of mixed quality characteristics based on the KPCA technique has been presented by Mashuri et al.⁶⁸. Recently, Lee et al.⁶⁹ presented new multivariate control charts by Hotelling's T^2 statistics and Q statistic based on KPCA approach for rapidly detecting a worn cutting tool and thus avoiding catastrophic tool failures products with unacceptable surface finish, and defective product. Their proposed method converts raw multi-sensor data into principal component space, and then, the KPCA-modified data are used to calculate T^2 and Q values to develop control charts.

2.3 Neural network (NN) and deep learning (DL)

Unlike linear models, the NN is based on a complex, divisional data model. It includes multiple layers to provide you with unique and precise output. However, the model is still based on linear regression but uses multiple hid-

den layers; this is why it is called a NN. In the paper of Arkat et al.⁷⁰, they designed a NN based model to forecast and construct residuals CUSUM chart for multivariate Auto-Regressive of order one, AR(1), processes. The comparison of the performance of the proposed method with the time series-based residuals chart and the auto-correlated MCUSUM chart was made. DL is a subset of ML, which is essentially a NN with multi layers. Recently, Lee et al.⁷¹ proposed a variational autoencoder (VAE) approach to monitor high-dimensional processes in the presence of non-linearity and non-normality assumptions. They demonstrated the effectiveness and applicability of the proposed VAE-based control charts in comparison with the existing latent variable-based charts through a simulation study and also via real data from a TFT-LCD manufacturing process. Chen and Yu⁷² suggested a novel recurrent neural network (RNN) residual chart with a DL technique to recognize mean shifts in autocorrelated processes. A comparison study with some typical methods such as special causes control chart and backpropagation network residual chart demonstrate that the RNN-based chart provides the best performance for monitoring mean shifts in autocorrelated manufacturing processes. The readers can find more reference about this technique based on control charts, for example, see Niaki and Abbasi⁷³, Chen et al.⁷⁴, and Diren et al.²⁸.

3 Machine Learning (ML) techniques based Control Chart Pattern Recognition (CCPR)

Entering the 21st century, the world has changed dramatically with the development of information technology, this is the beginning of the era of big data. This comes with a marked increase in the general interest in ML. The interpretation of control charts is mainly based on rough rules (i.e. heuristics) which depend greatly on the experience and judgment of the operator. It is therefore very important to make sure that they are well trained. Consequently, expert systems were born and developed in the industry. An expert system is software that is linked to at least two data sources: a database that contains a set of rules and a data flow that comes from the process to be controlled. The rules are based on the knowledge of experts in the field and are encoded as logical conditions. Everything is connected to a motor inference that applies the rules. The latter produces a result that is then communicated to users through a graphical interface and is used as a decision support tool. More precisely, an expert system is a software capable of answering questions, by reasoning from known facts and rules. However, the period of popularity of expert systems is relatively short, from the end of the 1980s, NNs are beginning to be used to automate the reading and interpretation of control (see Pugh⁷⁵). Since that time, pattern recognition, in general, is dominated by ML, is widely developed. There are several moti-

variations for using ML algorithms for CCPR purposes. The first and probably the main motivation is that several researchers demonstrated that the ML-based CCPR model outperforms their alternative models in many practical situations. For example, Li et al.⁷⁶ proposed a SVM-based CCPR framework and demonstrates that this model can accurately classify the source(s) of out-of-control signal and even outperforms the conventional multivariate control scheme. There also other motivations for applying ML-based CCPR models. For example, Guh⁷⁷ stated that the NN models are capable of learning and self-organizing and hence are useful in pattern recognition and can recall patterns learned from noisy or incomplete representations which are practically impossible to detect by operators, even with the assistance of an expert system. This makes the ML-based approaches suitable for CCPR because CCPs are generally contaminated by common cause variations. In addition, Diren et al.²⁸ reported that traditional CCPR models are not able to predict unexpected new situations while ML techniques that can effectively predict the unexpected new situations by learning from the historical data. This section reviews some important references about the most popular ML algorithms used in recognition of patterns on control charts including classification and regression tree (CART), decision trees (DTs), SVMs, NNs, and DL.

3.1 Regression tree (CART) and Decision tree (DT) based CCPR

A DT is a decision support tool representing a set of choices in the graphic form of a tree. Geometrically, construct a decision tree decision is to partition the space of data attributes in areas where each region represents a class. During prediction, when data is in this region then the decision tree assigns it the corresponding class. In literature, there are different methods to construct one or more decision trees from a learning data set. The common goal of each method is to determine the optimal test sequence to partition the space of attributes into homogeneous regions. Very recently, Zaman and Hassan⁷⁸ demonstrate the development of fuzzy heuristics and the CART technique for CCPR and compare their classification performance. The results show the heuristics Mamdani fuzzy classifier performed well in classification accuracy (95.76%) but slightly lower compared to the CART classifier (98.58%). This study opens opportunities for deeper investigation and provides a useful revisit to promote more studies into explainable AI.

3.2 Neural network (NN) and deep learning (DL) based CCPR

In the paper of Hachicha and Ghorbel⁷⁹, a survey of CCPR literature, the majority of the reviewed articles use the NN approach. It is reported that for the period 1988 to 2000, 9% of the revised publications use NNs and that

for the period 2001 to 2010, that number climbed to 25%. This trend then accelerates for the period from 2010 to 2021 (see Fig. 2 and Fig. 3). This observation is supported by the number of articles published on NN which shows an average annual increase of 10-15% for this period (source from Web of Science). Pugh⁷⁵ was the first author to experiment with NN and control charts. He concludes that NN is as effective as traditional control charts for detect changes in average values following a surge (by comparing the ARL) and NN was found to perform reasonably well under most conditions. This study constitutes the proof of concept of NN in CCPR. Pham and Oztemel⁸⁰ were the firsts described the structures of pattern recognition systems which made up of independent multi-layer perception. They found that these composite pattern recognition systems have better classification capabilities than their individual modules. Cheng⁸¹ also concluded that hybrid networks are more efficient than networks singular. Addeh et al.⁸² proposed a CCPR procedure based on optimized radial basis function neural network (RBFNN). The proposed method consists of four main modules: feature extraction, feature selection, classification and learning algorithm. In addition traditional patterns that have considered in literature including the normal, cyclic, increasing trend, decreasing trend, upward shift and downward shift, they investigated the stratification and systematic patterns as well. They tested RBFNN-based CCPR model based on a dataset containing 1600 samples (200 samples from each pattern) and the results showed that the proposed method has a very good performance. Yu et al.⁸³ developed an effective and reliable DL method known as stacked denoising autoencoder (SDAE) for CCPR in manufacturing processes. Recently, Xu et al.⁸⁴ proposed an efficient one-dimensional Convolutional Neural Network (1D-CNN) to applied for CCPR purposes. They showed that their method achieves 98.96% average recognition accuracy after 30 repeated tests as well as has better generalization ability when there is an error between the estimated value and true value of mean or standard deviation, which are satisfactory results. Yang and Zhou⁸⁵ developed online CCPR systems using NN0 ensemble also neglecting how the correlation coefficient is biased when abnormal patterns occur, thus training one CCPR system for each of the studied autocorrelation levels. Fuqua and Razzaghi⁸⁶ proposed a cost-sensitive classification scheme within a deep convolutional neural network (CSCNN) to fill the literature gap of developing computationally-efficient methods of CCPR classification for large time-series datasets in the presence of imbalance. To show the benefits of the method, they conducted an extensive experimental study using both simulated and real-world datasets based on simple and complex abnormal patterns. For more information, see examples some publications as Pham and Wani⁸⁷, Yang and Zhou⁸⁵.

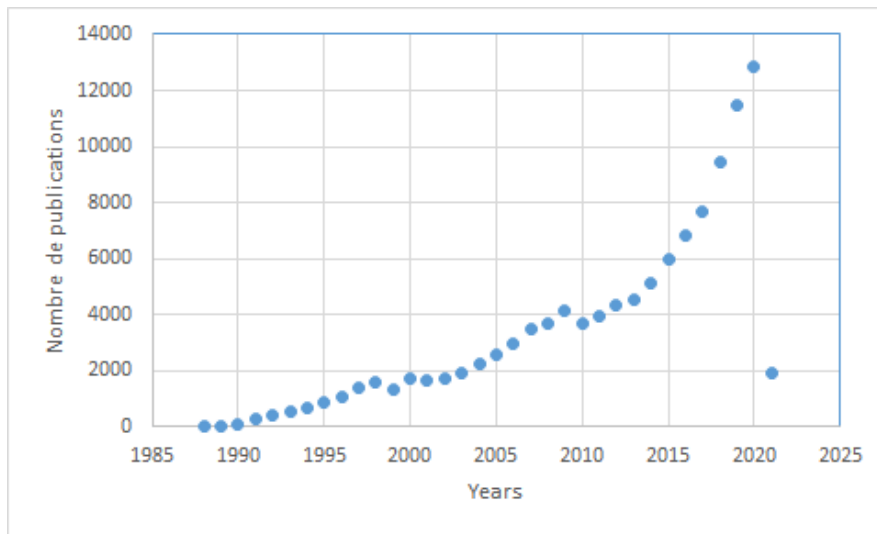


Figure 2: Number of publications on NN from 1988

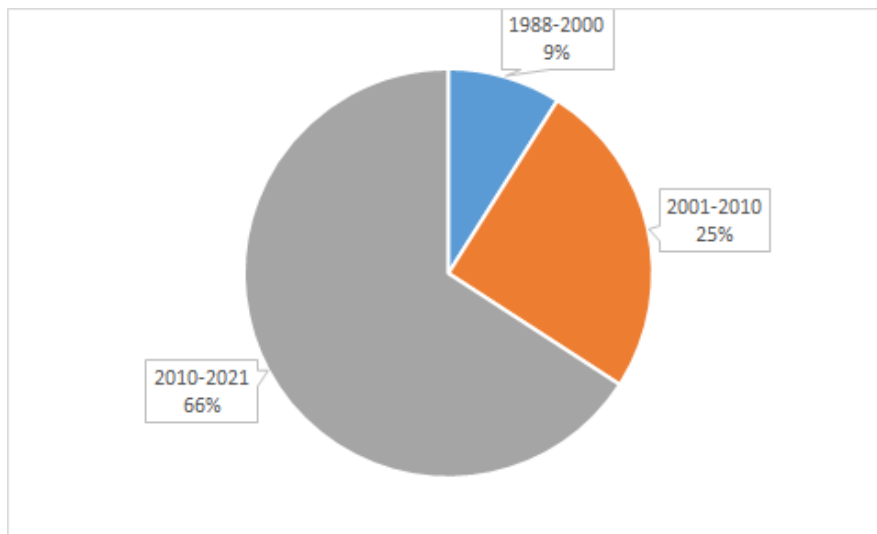


Figure 3: Percentage of number of publications on NN according to the period

3.3 Support vector machines (SVM)based CCPR

SVM are new statistical learning techniques proposed by V. Vapnik in 1995. They help to address diverse issues as classification, regression, fusion, etc. The essential idea of SVM consists in projecting the data of the input space (belonging to two different classes) non-linearly separable in a space of greater dimension called space of characteristics in such a way that the

data becomes linearly separable. In this space, the technique construction of the optimal hyperplane is used to calculate the function of classification separating the two classes (see Figure 4). In other words, the algorithm creates a line or a hyperplane which separates the data into classes.

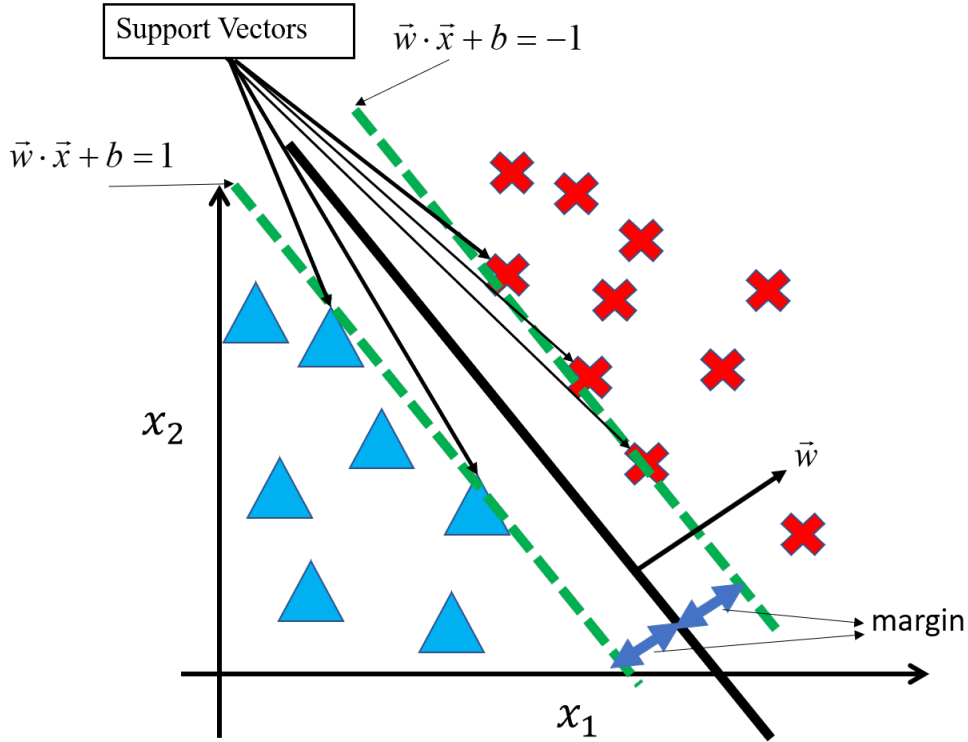


Figure 4: Principle of SVM techniques

In this subsection, we will summarize some recent applications and extensions of SVM for the CCPR case. Ranaee et al.⁸⁸ study a novel hybrid intelligent system that includes three main modules, in which two modules, SVM technique is used to searching for the best value of the parameters that tune its discriminant function (kernel parameter selection) and upstream by looking for the best subset of features that feed the classifier. Simulation results show that the proposed algorithm has very high recognition accuracy. A hybrid independent component analysis (ICA) and SVM is proposed for CCPR (Lu et al.⁸⁹), the results showed that is able to effectively recognize mixture control chart patterns and outperform the single SVM models, which did not use an ICA as a preprocessor. Lin et al.⁹⁰ presented a SVM-based CCPR model for the online real-time recognition of seven typical types of abnormal patterns, assuming that the process observations come from an AR(1) model. Through an extensive simulation study, they showed that the proposed SVM-based CCPR model can effectively on-line recognize unnatural patterns in both independent and autocorrelated processes.

In addition, they indicated that the new model has a better recognition accuracy and ARL performance than the existing learning vector quantization network CCPR model provided by Guh⁷⁷. Du et al.⁵⁹ integrated wavelet transform and improved particle swarm optimization-based support vector machine (P-SVM) for online recognition of concurrent CCPR. In other research, original SVM demonstrates poor performance when applied directly to these problems. Xanthopoulos and Razzaghi⁹¹ improve SVM by using weighted support vector machines (WSVM) for automated process monitoring and early fault diagnosis. They show the benefits of WSVM over traditional SVM, compare them under various fault scenarios. Readers can refer to many other references, see Wang⁹², Ranaee and Ebrahimzadeh⁹³, Lin et al.⁹⁰, Zhou et al.⁹⁴, la Torre Gutierrez and Pham⁹⁵.

4 Interpreting out-of-control signals using Machine Learning (ML)

When the manufacturing process has more than two characteristics for monitoring, it should be often solve with different univariate control charts. However, when these characteristics have a high correlation or sampling in a short duration, the false alarms may be appeared. Therefore, it is necessary to use multivariable control charts for monitoring quality problems. Hotelling's T^2 chart (Hotelling²⁵), MCUSUM chart (Woodall and Ncube²⁶), and MEWMA chart (Lowry et al.²⁷) are common multivariate charts were used in MSPC. However, a challenge of these traditional multivariate control charts is that they are just only able to provide the general mean shifts in vector, i.e., out-of-control signals of the process. It is impossible for these charts to indicate which variable(s) or a group of variables is responsible for out-of-control signals of the process. Numerous researchers have paid attention to the topic which to find a variable or a number of variables or a set of variables responsible for the signals when a multivariate process is in the out-of-control state. From the past decades, the idea integrating ML to multivariate control charts as an effectively approach. Recently, this approach seems more reasonable when the system of manufacturing has become more automatic. Thus, this section will give a look at the literature about ML methods for interpreting control charts in the multivariate process.

The first encouragement integrating ML methods to interpret signals of multivariate control charts in the quality control process has been discussed from the beginning of the 2000s with the publication of Wang and Chen¹⁹. Particularly, they used a neural-fuzzy model (a four-layer fully connected feed-forward network with a back-propagation training rule) for both detecting and classifying phases. An experiment for a bivariate process was conducted demonstrated that the proposed method reaches higher performance than the previous multivariate T^2 control chart. Lower out-

of-control ARLs and more classification accuracy results of the proposed method have been recorded. Then, Low et al.²⁰ continued highlight NNs as the method contribute to detect more anomaly patterns and more sensitive than traditional charts through out-of-control ARL and the numerous abnormal instants detected. Chen and Wang⁹⁶ suggested using a model of NNs, a three-layer fully connected feed-forward network with a back-propagation training rule, based multivariate χ^2 control chart to investigate cause variable(s) of signals of bivariate process. The significant advantages are showed that the model can indicate both responsible variables (s) and the magnitude of the shifts in case the multivariate χ^2 control chart has sudden shifts in the mean vector. S. T. A. Niaki and Abbasi²¹ suggested multilayer perceptron (MLP) network, a type of NNs, to classify patterns to explore variables or the set of variables that caused the fault of the process. The authors also make a comparison between MLP based Hotelling's T^2 multivariate Shewhart (MSCH) and based multivariate Shewhart (MS) chart, respectively. The results showed that the proposed MLP MSCH has a stronger performance. Cheng and Cheng⁹⁷ suggested to use 3-layer fully connected feed-forward network with a back-propagation training rule as an algorithm of NN for classifying out-of-control signals. The authors also recommend using SVMs which are considered as the method that has the same performance although it has more advantages than NN. On the contrary, Guh and Shiue⁹⁸ suggested using DT techniques instead of NNs based model to interpret which variable or group of variables has caused the out-of-control signals. They also demonstrated that the implementation of the DT approach gained results faster than 25 times than the NN one. According to Yu et al.⁹⁹, a selective NN ensemble approach named Discrete Particle Swarm Optimization (DPSOEN) algorithm has a significant performance to provide the source(s) of out-of-control. Alfaro et al.¹⁰⁰ proposed to use a multi-class exponential loss function (SAMME) algorithms, an extension of AdaBoost for classifying which variables have to responsible for the out-of-control signals. They showed that the proposed method has more significant performance than ones in the study of S. T. A. Niaki and Abbasi²¹. Verron et al.¹⁰¹ presented a Bayesian network-based control chart approach to detect and isolate fault variable(s) of a multivariate process. A DT learning-based model for bivariate process is recommended in a study of He et al.¹⁰² to identify the cause of faults. Cheng and Lee³¹ suggested using a SVM-based ensemble classification model for interpreting the out-of-control signal of a multivariate process by indicating the caused variable(s). An experiment comparison showed the significant performance of the proposed method in comparison with the single Support Vector Classification (SVC) model, bagging and, AdaBoost. Moreover, Carletti et al.¹⁰³ proposed Depth-based Isolation Forest Feature Importance (DIFFI) approach based Isolation Forest (IF) algorithm, the one from the idea as the DT to interpret the cause of faults in the process. The authors also make a comparison

with the Permutation-based Importance (PIMP) approach. Recently, Song et al.¹⁰⁴ recommend using a NN) method like instance-based Naive Bayes (INB) algorithm to classify which variables are the cause of out-of-control signals. This is well implemented for both small and large number variables. This also overcomes two disadvantages of previous studies as independence assumption and ignorance of the features of a test instance. Furthermore, very recently, Diren et al.²⁸ conduct a study with a variety of ML techniques including Naive Bayes-kernel (NB-k), K-Nearest Neighbor (KNN), DT, NN, Multi-Layer Perceptron (MLP), and DL to find the variables responsible for the out-of-control signals based types of faults. Performance comparison of these techniques is explored. Salehi et al.¹⁰⁵

5 Difficulties and Challenges for application of Machine Learning in statistical process control charts

It is evident that firms and corporations are rapidly getting smarter by adding intelligence into their process to drive continuous improvement, knowledge transfer, and intelligent decision-making procedures. This increases the demand for advanced AI and SPC tools and also effectual integrated techniques in various production stages to decrease the cost of production, improve overall productivity, improve product and process quality, reduce downtime, and etc. One of the most successful integrations is using ML algorithms, as an important subset of AI, in development, pattern recognition, and interpreting of control charts, as the main goals of SPC. To meet this need, several ML-based approaches have been developed by researchers and scientists that some of them are reviewed in the previous two sections. However, most of these tools have been introduced in laboratory environments and many difficulties and challenges still exist in their applications in practical environments. Implementation of an efficient ML algorithm that performs well in an industrial environment as well as produces reliable results is not very easy. Accordingly, it can be said that although ML is an efficient and widely-used technique for solving nowadays complex problems, like any other technique, it should be implemented as a solver due to its difficulties and challenges. Although data analysts may face a variety of challenges during the designing and implementation of ML algorithms in development, pattern recognition, and interpreting of control charts that we can not address them all here, however, in what follows, we will list some of them that are most appeared in daily operation problems.

5.1 Non-stationary Processes

Although several studies have been done for developing ML-based control charts, CCPR frameworks in the presence of autocorrelated observations

(see for example Lin et al.⁹⁰, Chen and Yu⁷², Kim et al.⁴⁴, and Yang and Zhou⁸⁵), most of these studies are based on the assumption that the process is stationary. However, most processes in the manufacturing industries are non-stationary in particular for complex industrial processes which in general show non-stationary process characteristics, revealing a time-varying mean and/or variance or even time-varying autocovariance (Zhao et al.¹⁰⁶). This phenomenon makes monitoring a complex task no matter the quality characteristic to be monitored is univariate or multivariate. Non-stationarity in processes' behavior frequently occurs due to several factors such as seasonal changes, processes that involve emptying and filling cycles, throughput changes, the presence of unmeasured disturbances, and also the nature of the process itself (Chen et al.¹⁰⁷). In these cases, interpreting out-of-control points is a challenge as studies on the topic almost always make assumptions about the distribution. Ketelaere et al.¹⁰⁸ presented examples of non-stationary processes from the industrial machinery monitoring context and agriculture industry. Another examples of non-stationary processes in industrial environments are discussed in Chen et al.¹⁰⁷ and Liu and Chen¹⁰⁹. Monitoring non-stationary processes have its challenges and difficulties and it has to be done carefully since there are many hidden problems. For example, it is difficult to detect the abnormal patterns of non-stationary observations because they may be hidden by the normal non-stationary variations (Zhao et al.¹⁰⁶). In addition, Lazariv and Schmid¹¹⁰ showed that for some processes and change-point models the ARL does not exist. This is a very important issue since the ARL is the most popular measure for the performance of control charts. In such situations, the traditional SPC techniques fail at monitoring such processes and it is important to have tools that can correctly detect changes in non-stationary processes (Lazariv and Schmid¹¹¹).

5.2 Big Data analysis

The term big data refers not only to the size or volume of data but also to the variety of data and the velocity of data. These features impose some challenging issues to the data analyst facing various big data monitoring problems. One of the main challenges for monitoring big data based on ML techniques is the training (Phase I) dataset that is expected to contain both in-control and out-of-control process observations (Qiu¹¹²). It is known that completing Phase I is critical to successful Phase II monitoring and has a strong influence on the performance and suitability of the ML algorithm to get accurate results and to avoid false predictions. However, in SPC applications, we usually have an in-control dataset only and there is no information about out-of-control situations in the training data. We know that it is very important to provide a training data set that entirely represents the structure of the problem. To tackle this deficiency, the idea of artificial contrasts

and one-class classification methods have been suggested by authors such as Tuv and Runger¹¹³ and Sun and Tsung⁴¹. Another challenge in monitoring high dimensional data sets is the fact that not all of the monitored variables are likely to shift at the same time, thus, some method is necessary to identify the process variables that have changed. In high dimensional data sets, the decomposition methods used with multivariate control charts can become very computationally expensive Reis and Gins¹¹⁴. To serve the purpose, many scientists proposed feature selection techniques to monitor subsets of potentially faulty variables instead of monitoring a sequence of whole variables to improve detection performance (see for example Capizzi and Masarotto¹¹⁵). However, in such cases, the key questions that have not to be answered yet are a) what kind(s) of features are appropriate to use for a specific big data monitoring problem, b) how many features should be extracted for process monitoring, and c) whether the original goals of process monitoring have been substantially compromised by using the selected features Qiu¹¹².

5.3 Monitoring image data

Thanks to the rapid developments of digital devices like sensors and computers and using them increasingly in industrial and medical applications, intelligent decision-making tools such as machine vision systems (MVS) has gradually taken the place of human-based inspections in many factories due to their ability to provide not only dimensional information but also information on product geometry, surface defects, surface finish, and other product and process characteristics Megahed et al.¹¹⁶. A MVS is a computer-based system for analyzing and processing image data that is provided by image-capturing devices (e.g., cameras, X-ray devices, or vision sensors). New studies show that implementing MVSs in industrial environments could be fully utilized to improve the quality of the product Zuo et al.¹¹⁷. In this regard, researchers developed a new interdisciplinary field of research by integrating MVS approaches and SPC principles. This new field applied SPC tools for monitoring the process quality using images. There are several applications in industrial and medical areas that image monitoring can be used to check the stability of the process state. For instance, monitoring the brightness of the cover in the printing process of a journal or monitoring the changes of tumors and vascular. Through an extensive review of image-based control charting methodologies, Megahed et al.¹¹⁶ emphasized that using MVS-based monitoring procedure is superior to visual inspection with respect to, (1) monitoring processes with high production rates; (2) performing multiple simultaneous tasks with different objects; (3) their ability to cover all the ranges of the electromagnetic spectrum, as in the use of magnetic resonance imaging (MRIs) and X-rays in medical applications; (4) the lack of susceptibility to fatigue and distraction; and (5) in some cases, the

use of MVSs is cheaper than the use of human inspectors and it is expected that the cost of MVSs will continue to decrease over time. However, there are several challenges in implanting image monitoring in practical situations. The first challenge is that the number of pixels in a simple cell phone image nowadays is around 4 million pixels and thus, we have to monitor a process with 4 million components over time that faces us to high-dimensional problems. Another challenge is that the neighboring pixels within an image are often highly correlated. This correlation can result in a considerable amount of data redundancy and ignoring the correlation can result in a high level of false alarms as well as poor performance once a fault occurs. In addition, there are several stages for successful impersonation of an image-based monitoring procedure such as the choice of the image-capturing device, the frequency of imaging, the set-up of the imaging to avoid lighting, alignment, the software to use for image analysis, the preliminary image processing, and the type of monitoring method to employ. There are no currently existing guidelines for guiding the practitioner through all of these decisions Megahed et al.¹¹⁶. Thus, the last challenge is providing easy- and clear-to-used guidelines to applied an efficient image monitoring model in practical applications.

6 Perspectives for Application of Machine Learning (ML) in statistical process control charts in Smart Manufacturing (SM)

Making processes smart and digitized, motivate researchers and scientists to develop effective ML strategies for anomaly detection in daily operations. For example, strategies to keep the production systems always dynamic in dealing with unexpected variations and abnormal patterns. Although recent studies have investigated new ML-based techniques in the development, pattern recognition, and interpreting of control charts in manufacturing, there still exists a significant potential for reducing the gap between the theory and application in modern industries. Addressing this gap will ensure that ML tools can be seamlessly integrated into factory operations. The following topics are recommended here for future research.

6.1 Auto-correlated processes and Non-stationary processes

Thanks to the rapid evolution of sensor technologies and the velocity of available data in modern industrial processes, a good ability has created to gather observations instantaneously that results in a high degree of autocorrelation within observations. In fact, the real-world data are in most cases auto-correlated. To deal with such data, most of the existing approaches are not

sufficient, and there is an essential need to develop new powerful ML tools to analyze these kinds of datasets. While the effect of autocorrelation on traditional SPC tools has been investigated by several authors, see for example Maragah and Woodall¹¹⁸ and Alwan⁹, a review paper by Apsemidis et al.³⁵ shows that a few studies have yet been conducted in the area of kernel-based ML methods for such data. In another comprehensive review paper, Weese et al.³⁴ recommended that although some few ML-based monitoring procedures have been proposed for autocorrelated data considering known time series model by researchers like Arkat et al.¹¹⁹, Issam and Mohamed⁵⁸, and Kim et al.¹²⁰, there is ample potential to develop new algorithms when the time series model is unknown. In literature of ML techniques based control chart pattern recognition there are also few investigations (Cuentas et al.¹²¹) that recognize relatively simple patterns such as process mean shift (Chinnam and Kumar¹²², Hsu et al.¹²³, and Hsu et al.¹²⁴) and process variance shift (Chinnam³⁷), while more complex patterns including trend, cycle, and systematic patterns only considered in the study of Lin et al.⁹⁰. Concurrently, the existing body of researches needs to be enhanced and improved the existing techniques of monitoring auto-correlated processes. In addition, there are several real-world examples that not only the observations are auto-correlated, but also they have non-stationary behaviors in which they are not oscillating around a common mean or its variance and autocovariance are changing over time. This phenomenon may happen due to several reasons. Researchers such as Ketelaere et al.¹⁰⁸, Chen et al.¹⁰⁷ and Chen et al.¹⁰⁷ presented examples of non-stationary processes in industrial environments. Up to now, most existing researches have focus on developing ML-based control charts and CCPR techniques for monitoring stationary processes and there is a remarkable need for developing such tools for handling time series data from non-stationary processes. This gap should be filled by new researches. Recently, Tran et al.¹²⁵ and Nguyen et al.¹²⁶ proposed Long Short Term Memory networks (LSTM) and LSTM Autoencoder techniques for monitoring multivariate time series data from non-stationary processes. These techniques can be also employed as efficient solutions for CCPR problems involving auto-correlated non-stationary data for future study. Section 7 of the current chapter provides a good discussion for bearing failure anomaly detection based on the LSTM method. Finally, Explainable AI techniques (see Rudin¹²⁷) should also be used to develop frameworks for interpretation of out-of-control points in this context.

6.2 Big data analysis

Nowadays in smart factories, a wide range of sensors have embedded in several devices of production lines as well as are connected to many computers for data analysis, management, and visualization, which brings fruitful business results in the long run. These advanced technologies created the

concepts of high-volume, high-dimensional, and high-velocity data that are called in brief Big data. This type of data always has complex natures as well as often has hierarchical or nested structures. In such situations, traditional SCP methods are incapable of monitoring such data and existing methodologies should be stretched to new limits. Although data-driven ML algorithms have a good potential to do this end, there is poor literature about ML-based studies for anomaly detection and pattern recognition using data sets that would be considered big data (Wang and Jiang¹²⁸, Jin et al.¹²⁹, Qiu¹³⁰, and Sparks and Chakraborti¹³¹). Thus, there is a tremendous opportunity and significant need for the development of advanced ML tools for both anomaly detection and CCPR. For example, Qiu¹¹² provided a comprehensive discussion on some recent SPC methods in the presence of big data and recommended the following research directions as further research. (i) feature selection methods have been suggested by some authors to simplify the computations for monitoring big data sets. In such cases, the key questions that have not been properly addressed yet in the literature and future research is needed in this direction are a) what kind(s) of features are appropriate to use for a specific big data monitoring problem, b) how many features should be extracted for process monitoring, and c) whether the original goals of process monitoring have been substantially compromised by using the selected features. (ii) process observations in SPC literature are widely assumed to be either independent or following some specific parametric time series models such as ARMA models. These assumptions are rarely satisfied in practice, especially when we are dealing with big data sets. More precisely, in the context of big data, process observations have many other complicated structures. Thus, developing SPC tools to properly accommodate such data structures will be an attractive research area. (iii) In practice, the performance of a process is often affected by various covariates that can provide some helpful information to us. Therefore, taking this information into account in developing and designing new SPC tools can improve the efficiency of the monitoring procedures. However, there is no study in the SPC literature yet regarding the proper use of helpful information in covariates. All these topics can also be investigated based on ML methods for both anomaly detection, interpreting out-of-control signals, and CCPR as well. More discussions on this topics can be found in Megahed and Jones-Farmer¹³² and Reis and Gins¹¹⁴.

6.3 Real word implementation and hyperparameters optimization

Scientists and engineers believe that we are at the beginning of the fourth industrial revolution that is being called Industry 4.0. Broadly speaking, this revolution has been happening by decreasing human intervention and adding intelligence into the production processes and service operations. Digitaliza-

tion and computerization enable manufacturers/managers to make their own smart factories/companies as a unified digital ecosystem of all the different works aspects using advanced technologies to organize and optimize their production/service cycles. In this situation, companies and corporates have begun to adapt and implement the state-of-the-art into daily operations to improve production efficiency, flexibility, and reduce cost (See for example Malaca et al.¹³³). On the other hand, Woodall and Montgomery¹³⁴ express that “Despite the large number of papers on this topic we have not seen much practical impact on SPC”. This is an unacceptable face of SPC literature that may occur because of several reasons. Weese et al.³⁴ also stated that there are very few discussions in the related literature that, i) address the step-by-step procedures of selection and operation of algorithm in practical situations, ii) conduct an illustrative Phase I analysis, and iii) provide some advice on how to apply the methods in practice, including how to establish an in-control training sample or how large training data size is needed to algorithm learned effectively. As an example of concerns (i) and (iii), one of the most important stages of ML-based algorithms implementation is hyperparameter optimization which is also known as hyperparameter tuning. Hyperparameters are those that lead to the highest accuracy and/or least error in the validation set and provide the best results for the problem they are solving. It is important to note that the hyperparameter is different from the model parameter. Hyperparameters are the model arguments that should be determined before the learning process begins and they are not learned from the training data like model parameters. For example, K in KNN, kernel type and constants in SVMs, number of layers, and neurons in neural networks are some of the well-known hyperparameters. These hyperparameters can be determined by maximizing (e.g. the accuracy) or minimizing (e.g the loss function) the specified metrics. Although these hyperparameters play a crucial role in utilizing ML algorithms that the effectiveness of the algorithm largely depends on selecting good hyperparameter values, surprisingly, most of the studies applying ML in SPC have not considered hyperparameter optimization in their studies. Bochinski et al.¹³⁵ proposed an evolutionary algorithm-based hyper-parameter optimization approach for committees of multiple CNNs. Trittenbach et al.¹³⁶ developed a principled approach using the Local Active Min-Max Alignment method to estimate SVDD hyperparameter values by active learning. In a ML-based SPC investigation, Trinh et al.¹³⁷ investigated the application of one-class SVM to detect anomalies in wireless sensor networks with data-driven hyperparameter optimization. Also, Wu et al.¹³⁸ proposed an effective technique for Hyperparameter tuning using reinforcement learning. Based on the above-mentioned discussions, there is a large gap between the theories and assumptions in literature and real demands in industrial environments that should be reduced through future research. Accordingly, it is recommended to scientists for providing illustrative guidelines for probably non-specialist practitioners to show them

clearly how to implement the method in their problems. Moreover, it is also important to prepare the source code of test designs using ML because most of them have no explicit expression of control limits and ARL. This would make the implantation of the proposed methods easy for practitioners.

6.4 Integration of SVM and NN techniques

It is known that both SVM and NN are powerful ML algorithms in the anomaly detection and pattern recognition contexts because of their impressive results which are reported in many references. However, each of them has its advantages and disadvantage. For example, the structural risk minimization of SVMs benefits their performance, in contrast with the empirical risk minimization of NNs, which creates problems. While NNs try to minimize the training error, the SVMs minimize an upper bound of the error, something that enables them to generalize easier even when the dataset is small. Furthermore, SVMs find a global solution and cannot be stuck in local minima, in contrast with the NNs (Apsemidis et al.³⁵). So, their combination may lead to aggregation of benefits to serve as a unified attractive tool for ML-based SPC activities. For example in an anomaly detection problem, one might utilize a deep NN and have the final classification via SVM at the output layer. It is likely to have better classification results compared to ordinary NN. Recently, Hosseini and Zade¹³⁹ suggested a new hybrid technique called the MGA-SVM-HGS-PSO-NN model for detection of a malicious attack on computer networks by combining SVM and NN techniques. Their proposed method includes two stages, a feature selection stage and an attack detection stage. The feature selection process was performed using SVM and a Genetic Algorithm (GA). On the other side, the attack detection process was performed using an NN approach. The performance of the MGA-SVM-HGS-PSO-NN method was compared with other popular techniques such as Chi-SVM, NN based on gradient descent and decision tree, and NN based on GA based on performance metrics like classification accuracy, training time, the number of selected features, and testing time on the basis of the well-known NSL-KDD dataset. They showed that the proposed method is the best performing method on all criteria. For example, the proposed MGA-SVM-HGS-PSO-NN method can attain a maximum detection accuracy of 99.3%, dimension reduction of NSL-KDD from 42 to 4 features, and needs only 3 seconds as maximum training time. In a good review paper on ML Kernel Methods in SPC, Apsemidis et al.³⁵ showed that while 43% of the papers compare the SVM and NN algorithms and in 51.9% there is no reference of NN in the SVM method, only 5.1% of the cases the SVM and NN are combined to work together in the proposed method. Thus, there is a large room here for developing new methods and improving existing ML-based anomaly detection and CCPR models. For instance, investigating the possible design of control charts for monitoring

stationary and non-stationary multivariate time series data with LSTM or Autoencoder CNN combined with SVDD technique can be considered as a good research topic (see Tran et al.¹²⁵ and Nguyen et al.¹²⁶).

6.5 ML algorithm in the presence of drift

One of the assumptions in supervised learning is that the mapping function f is assumed to be static, meaning that it does not change over time. However, in some problems, but not all problems, this assumption may not hold true. It means that the structure of data can change over time and hence the relation between input and output would be dynamic. This phenomenon in the ML literature known as concept drift and may happen due to several reasons. Ignoring concept drift while we are selecting and learning the predictive model can affect the prediction power of the algorithm. To tackle this problem, many adaptive learning techniques have been proposed by researchers like Žliobaitė et al.¹⁴⁰ and Gama et al.¹⁴¹. However, to the best of our knowledge, there is no study for ML-based control charts, pattern recognition, and interpreting out-of-control signals by considering the concept drift. So, there is a large potential here for more researches.

6.6 Data fusion and feature fusion

Data fusion and features are newly developed fields in data science that deal with the problem of the integration of data and knowledge from multiple sources and reducing the features' space of raw data, respectively. This technique can improve available information of data in the sense of decreasing the associated cost, increasing the data quality and veracity, gathering more related information, and increasing the accuracy of ML-based tools. Especially, it can be useful in smart factories with multisensor environments. For example, the main advantages of data fusion are discussed in more detail in the biosurveillance area by Shmueli and Fienberg¹⁴² (pp. 123-133). Castanedo¹⁴³ classified the data fusion techniques into three main categories as ,i) data association, ii) state estimation, and iii) decision fusion. In addition, feature fusion techniques can improve the ability of mixture CCPRs. Recently, Zhang et al.¹⁴⁴ proposed a CCPR model based on fusion feature reduction (FFR), which makes the features more effective, and fire-works algorithm-optimized MSVM. They showed that the proposed method can significantly improve the recognition accuracy and the recognition rate and the run time of CCPR as well as deliver satisfying prediction results even with relatively small-sized training samples. Another CCPR technique based on the feature fusion approach is presented in Zhang et al.¹⁴⁵. In conclusion, developing new and refining existing ML-based control charts and ML-based CCPR models, as well as interpreting out-of-control techniques based on data fusion and feature fusion methods are good directions for

future research of the scientist in the field of this chapter (Weese et al.³⁴).

6.7 Control chart for complex data types

Data in the smart factories are nowadays collected with a high frequency, high dimension, complex structure, and large variety which cannot be treated straightforwardly. These new circumstances create the concept of complex data. Functional data, compositional data, and topological data are some important types of complex data. To handle such data, new data analysis methods have developed or the existing techniques have refined by some researchers. For example, Topological Data Analysis (TDA) has proposed to analyze topological data that emerges as a powerful tool to extract insights from high-dimensional data. The core idea of TDA is to find the shape, the underlying structure of shapes, or relevant low dimensional features of complex structure and massive data. In Umeda et al.¹⁴⁶, the application of TDA is used to describe the time-series DL for analyzing time series data and anomaly detection. In particular, two key technologies- Mapper and persistent homology are applied in both supervised learning and unsupervised learning. Mapper presents the distinguishing features of a set of data as an easy-to-understand graph. Persistent homology is a technology that numerically captures a data shape in detail. This paper developed an anomaly detection technology for time-series to detect an abnormal state using TDA. Besides that, the data is becoming more and more related to functional data. The studies on monitoring functional data have drawn a lot of attention Colosimo and Pacella¹⁴⁷, Liu et al.¹⁴⁸, and Flores et al.¹⁴⁹. Anomaly detection methods for functional data based on functional PCA Yu et al.¹⁵⁰, wavelet functional PCA Liu et al.¹⁴⁸ are developed. However, the application of advanced ML on this type of data for development, pattern recognition, and interpreting of control charts still needs to be discovered. Thus, more efforts are needed to develop tests that use ML to track these types of data, need to find more documentation on ML methods suitable for them in order to write them correctly. For instance, although these studies have eliminated a lot of assumptions about the distribution of data when designing control charts with ML techniques, there are still independent data assumptions that do not exist in the data environment collected from IoT sensors. In general, there are still very few studies on this promising approach and further researches needs to be carried out to discover its numerous applications to the smart factory. Accordingly, developing advanced ML techniques to eliminate most of the assumptions of traditional SPC in development, pattern recognition, and interpreting of control charts for monitoring complex data types such as multivariate time series data, image data, and Big Data with complex structures is a high-potential area to carry out more researches. This will be a promising research direction to solve the problem of smart factory SPC implementation with Big Data.

6.8 Monitoring image data

Although applications of MVSs in industrial and medical shop floors have been increased dramatically and a huge number of possible applications exist here, but there are only a few papers dealing with image monitoring. Megahed et al.¹¹⁶ reviewed image-based control charts including univariate, multivariate, profile, spatial, multivariate image analysis, and medical image devices charts and addressed the capability of image-based monitoring in a much wider variety of quality characteristics. They noted that the use of image-based control charts differs from traditional applications of control charts in the SPC area. These differences can be attributed to a number of factors, which include the type of data being monitored, the rationale behind using control charts, and how the control charts are applied. Additionally, preprocessing of image data can also become a factor with 100% inspection since the data preprocessing time can be longer than the production cycle time. Therefore, these factors need to be considered when developing the control charting strategy. He et al.¹⁵¹ proposed a multivariate control charting method for both single and multiple faults. In their method, each image is divided into non-overlapping regions of equal size, and the mean intensities of these regions are monitored with a multivariate GLR-based statistic. Later, by extending the results of He et al.¹⁵², Stankus and Castillo-Villar¹⁵³ developed a multivariate generalized likelihood ratio control chart to identify process shifts and locate defects on artifacts by converting 3D point cloud data to a 2D image. They considered the surface dent in addition to two ordinary types of defects, surface curvature, and surface scratch, that does not identify by the existing methodologies. By means of a comparative study, Stankus and Castillo-Villar¹⁵³ showed that the new methodology has a significantly shorter out-of-control *ARL* than the He et al.¹⁵² methodology for the scratch and no significant difference in out-of-control *ARL* for the incorrect surface curvature. Zuo et al.¹¹⁷ reported that the existing research in the image-based SPC area has focused on either identification of fault size and/or location or detection of fault occurrence and there is limited research on both fault detection and identification. To handle such situations, they proposed an EWMA and region growing based control chart for monitoring of 8-bit grayscale images of industrial products. The results of the simulation study showed that the new method is not only effective in quick detection of the fault but also accurate in estimating the fault size and location. Recently, Okhrin et al.¹⁵⁴ provided an overview of recent developments on monitoring image processes. While we review some existing literature in this field, there are still some research opportunities in the integration of ML-based control charting methods and pattern recognition models with image data. It is known that with smart manufacturing, the amount of images collected from production lines is very big and each image may include millions of pixels that need ML approaches

to develop new control charts and CCPR frameworks. Many authors assume an independent residual process, while there is a natural spatial correlation structure of the pixels in neighborhoods. Therefore, there is a consequent need for some ML-based approaches for the successful monitoring of image processes. The existing methods, for instance, Okhrin et al.¹⁵⁵ and Yuan and Lin¹⁵⁶, can be improved to developing CNN and Transformers control charts to monitoring images in SM.

7 A Case Study: Monitoring and early fault detection in bearing

In this section, we present an application of ML based control chart for monitoring and early fault detection in bearing. AD in vibration signals is an important technique for monitoring, early detection of the failure, and fault diagnosis for rotating machinery. Very recently, Tran et al.¹²⁵, Tran et al.¹⁵⁷ and, Nguyen et al.¹²⁶ have developed very efficient methods with Long Short Term Memory networks (LSTM) and LSTM Autoencoder techniques in detecting anomalies for multivariate time series data. In this case study, we will combine both of these methods to propose a new ML based control chart that performs anomaly detection in an industry context. According to Nguyen et al.¹²⁶, we suppose that the autoencoder LSTM has been trained from a normal sequence $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, where N is the number of samples and $\mathbf{x}_t = \{x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(k)}\}, t = 1, 2, \dots$ is the value of the multivariate time series at the time t with k number of variables (these notations are from previous section). Using a sliding window of size m , the trained autoencoder LSTM can read the input sequence $\mathbf{X}_i = \mathbf{x}_t, \dots, \mathbf{x}_{t-m+1}$, encode it and recreate it in the output $\hat{\mathbf{X}}_i = (\hat{\mathbf{x}}_t, \dots, \hat{\mathbf{x}}_{t-m+1})$, with $i = m + 1, \dots, N$. Since these values has been observed from the data, one can calculate the prediction error $e_i = \|\hat{\mathbf{X}}_i - \mathbf{X}_i\|$, $i = m + 1, \dots, N$. The anomaly detection is then based on these prediction errors. The anomaly scores distribution of the training dataset is shown in Figure 5. In many studies, these error vectors are supposed that follow a Gaussian distribution and then used the maximum likelihood estimation method to estimate the parameters of this distribution. However, one can argue that the assumption of Gaussian distribution for error vectors may not be true in practice. To overcome the disadvantage of this method, Tran et al.¹²⁵ proposed used the kernel quantile estimation (KQE) control chart (Sheather and Marron¹⁵⁸) to automatically determines a threshold for time series anomaly detection. In particular, at the new time t , if $e_t > \tau$, x_t is classified as anomaly point and vice versa, see Tran et al.¹²⁵ for more details.

The experimental data were generated from a bearing test rig that was able to produce run-to-failure data. These data were downloaded from the

Prognostics Center of Excellence (PCoE) through a prognostic data repository contributed by Intelligent Maintenance System (IMS), University of Cincinnati (Qiu et al. ¹⁵⁹). According to (Qiu et al. ¹⁵⁹), vibrations signals were collected every 10 minutes with a data sampling rate was 20kHz and the data length was 20 480 sensor data points.

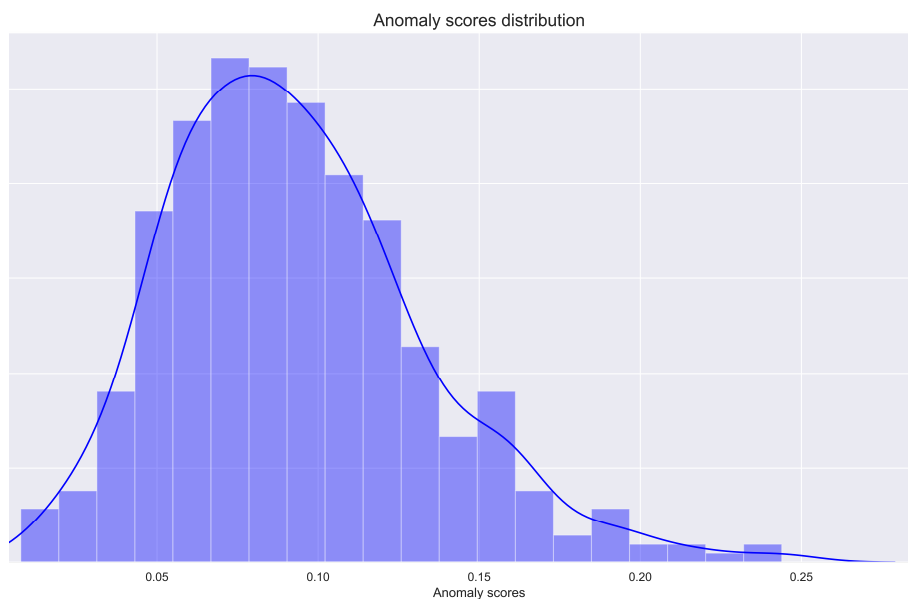


Figure 5: Anomaly scores distribution of the training dataset

This ML-based control chart allows for conditional monitoring and prediction of the upcoming bearing malfunction well in advance of the actual physical failure. It allows to automatically define a threshold value for flagging anomalies while avoiding too many false positives during normal operating conditions. The early detection of bearing failure is shown in the Figure 6, the bearing failure is confirmed at the end of this experiment (Qiu et al. ¹⁵⁹). This promising approach could provide a perfect tool to enable predictive maintenance implementation in SM.

8 Conclusion

Along with the development of technologies and AI, leading to production systems become more complex and modern-day by day. Therefore, the application of ML to SPC is an interesting and necessary trend that has been strongly developed in recent years to meet the needs of SM. In this chapter,

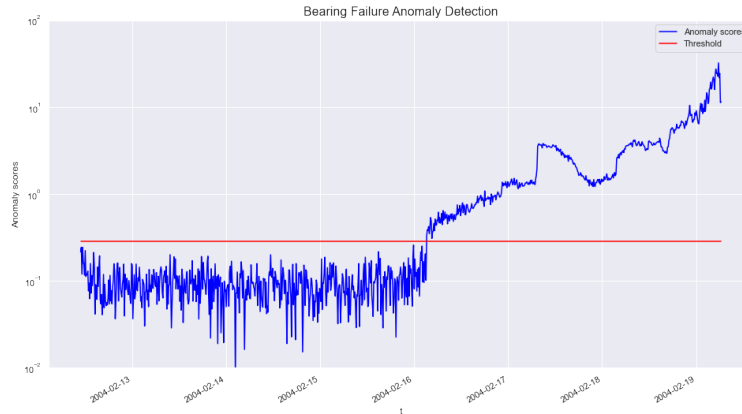


Figure 6: Bearing Failure Anomaly Detection

we have introduced different applications of ML in control chart implementation including designing, recognition trend, and interpreting. A literature review about these issues is discussed. Although there have been many achievements in research in this field, there are still many difficulties and problems that need to be solved in order to be able to apply control charts to SM. There still exists a significant potential for reducing the gap between theory and application in modern industries. A case study is also provided to present a ML-based control chart for monitoring and early fault detection in bearing.

References

- [1] F. Kadri, F. Harrou, S. Chaabane, Y. Sun, and C. Tahon. Seasonal arma-based spc charts for anomaly detection: Application to emergency department systems. *Neurocomputing*, 173:2102–2114, 2016.
- [2] G. Münz and G. Carle. Application of forecasting techniques and control charts for traffic anomaly detection. In *Proceedings of the 19th ITC Specialist Seminar on Network Usage and Traffic*, 2008.
- [3] P. H. Tran, K. P. Tran, T. H. Truong, C. Heuchenne, H. Tran, and T. M. H. Le. Real time data-driven approaches for credit card fraud detection. In *Proceedings of the 2018 international conference on e-business and applications*, pages 6–9, 2018.
- [4] P. H. Tran, C. Heuchenne, H. D. Nguyen, and H. Marie. Monitoring coefficient of variation using one-sided run rules control charts in the

- presence of measurement errors. *Journal of Applied Statistics*, pages 1–27, 2020.
- [5] P. H. Tran and C. Heuchenne. Monitoring the coefficient of variation using variable sampling interval cusum control charts. *Journal of Statistical Computation and Simulation*, 91(3):501–521, 2021.
- [6] V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3):1–58, 2009.
- [7] F. Y. Edgeworth. Xli. on discordant observations. *The london, edinburgh, and dublin philosophical magazine and journal of science*, 23(143):364–375, 1887.
- [8] W. A. Shewhart. Some applications of statistical methods to the analysis of physical and engineering data. *Bell System Technical Journal*, 3(1):43–87, 1924.
- [9] L. C. Alwan. Effects of autocorrelation on control chart performance. *Communications in statistics-Theory and Methods*, 21(4):1025–1049, 1992.
- [10] R. Noorossana and S. J. M. Vaghefi. Effect of autocorrelation on performance of the mcusum control chart. *Quality and Reliability Engineering International*, 22(2):191–197, 2006.
- [11] A.F.B. Costa and P. Castagliola. Effect of measurement error and autocorrelation on the \bar{X} chart. *Journal of Applied Statistics*, 38(4):661–673, 2011.
- [12] R. C. Leoni, A. F. B. Costa, and M. A. G. Machado. The effect of the autocorrelation on the performance of the t2 chart. *European Journal of Operational Research*, 247(1):155–165, 2015.
- [13] E. Vanhatalo and M. Kulahci. The effect of autocorrelation on the hotelling t2 control chart. *Quality and Reliability Engineering International*, 31(8):1779–1796, 2015.
- [14] R. S. Guh and Y. C. Hsieh. A neural network based model for abnormal pattern recognition of control charts. *Computers & Industrial Engineering*, 36(1):97–108, 1999.
- [15] J. A. Swift and J. H. Mize. Out-of-control pattern recognition and analysis for quality control charts using lisp-based systems. *Computers & Industrial Engineering*, 28(1):81–91, 1995.
- [16] Y. Guo and K. J. Dooley. Identification of change structure in statistical process control. *The International Journal of Production Research*, 30(7):1655–1669, 1992.

- [17] Z. Miao and M. Yang. Control chart pattern recognition based on convolution neural network. In *Smart Innovations in Communication and Computational Sciences*, pages 97–104. Springer, 2019.
- [18] T. Zan, Z. Liu, H. Wang, M. Wang, and X. Gao. Control chart pattern recognition using the convolutional neural network. *Journal of Intelligent Manufacturing*, 31(3):703–716, 2020.
- [19] T. Y. Wang and L. H. Chen. Mean shifts detection and classification in multivariate process: a neural-fuzzy approach. *Journal of intelligent manufacturing*, 13(3):211–221, 2002.
- [20] C. Low, C. M. Hsu, and F. J. Yu. Analysis of variations in a multivariate process using neural networks. *The International Journal of Advanced Manufacturing Technology*, 22(11):911–921, 2003.
- [21] B. S. T. A. Niaki and Abbasi. Fault diagnosis in multivariate control charts using artificial neural networks. *Quality and reliability engineering international*, 21(8):825–840, 2005.
- [22] E Western. Statistical quality control handbook. western electric co, 1956.
- [23] J. A. Swift. *Development of a knowledge based expert system for control chart pattern recognition and analysis*. PhD thesis, Oklahoma State University, 1987.
- [24] M. Shewhart. Interpreting statistical process control (SPC) charts using machine learning and expert system techniques. In *Proceedings of the IEEE 1992 National Aerospace and Electronics Conference@ m_NAECON 1992*, pages 1001–1006. IEEE, 1992.
- [25] H. Hotelling. Multivariate quality control. *Techniques of statistical analysis*, 1947.
- [26] W. H. Woodall and M. M. Ncube. Multivariate cusum quality-control procedures. *Technometrics*, 27(3):285–292, 1985.
- [27] C. A. Lowry, W. H. Woodall, C. W. Champ, and S. E. Rigdon. A multivariate exponentially weighted moving average control chart. *Technometrics*, 34(1):46–53, 1992.
- [28] D. Demircioglu Diren, S. Boran, and I. Cil. Integration of machine learning techniques and control charts in multivariate processes. *Scientia Iranica*, 27(6):3233–3241, 2020.
- [29] R. S. Guh and J.D.T Tannock. Recognition of control chart concurrent patterns using a neural network approach. *International Journal of Production Research*, 37(8):1743–1765, 1999.

- [30] K.L. Wu and M. S. Yang. A fuzzy-soft learning vector quantization. *Neurocomputing*, 55(3-4):681–697, 2003.
- [31] C. S. Cheng and H. T Lee. Diagnosing the variance shifts signal in multivariate process control using ensemble classifiers. *Journal of the Chinese Institute of Engineers*, 39(1):64–73, 2016.
- [32] Z. Kang, C. Catal, and B. Tekinerdogan. Machine learning applications in production lines: A systematic literature review. *Computers & Industrial Engineering*, 149:106773, 2020.
- [33] P. Qiu and X. Xie. Transparent sequential learning for statistical process control of serially correlated data. *Technometrics*, (just-accepted): 1–29, 2021.
- [34] M. Weese, W. Martinez, F.M. Megahed, and L.A. Jones-Farmer. Statistical learning methods applied to process monitoring: An overview and perspective. *Journal of Quality Technology*, 48(1):4–24, 2016.
- [35] A. Apsemidis, S. Psarakis, and J.M. Moguerza. A review of machine learning kernel methods in statistical process monitoring. *Computers & Industrial Engineering*, 142:106376, 2020.
- [36] M. Mashuri, H. Haryono, and M. Ahsan D.F. Aksioma, and W. Wibawati, and H. Khusna. Tr (r2) control charts based on kernel density estimation for monitoring multivariate variability process. *Cogent Engineering*, 6(1):1665949, 2019.
- [37] R.B. Chinnam. Support vector machines for recognizing shifts in correlated and other manufacturing processes. *International Journal of Production Research*, 40(17):4449–4466, 2002.
- [38] E. Byvatov, J. Sadowski U. Fechner, and G. Schneider. Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. *Journal of chemical information and computer sciences*, 43(6), 2003.
- [39] L. Li and H. Jia. On fault identification of mewma control charts using support vector machine models. In *International Asia Conference on Industrial Engineering and Management Innovation (IEMI2012) Proceedings*, pages 723–730. Springer, 2013.
- [40] F. Camci and R.B. Chinnam. General support vector representation machine for one-class classification of non-stationary classes. *Pattern Recognition*, 41(10):3021–3034, 2008.
- [41] R. Sun and F. Tsung. A kernel-distance-based multivariate control chart using support vector methods. *International Journal of Production Research*, 41(13):2975–2989, 2003.

- [42] X. Ning and F. Tsung. Improved design of kernel distance-based charts using support vector methods. *IIE Transactions*, 45(4):464–476, 2013.
- [43] T. Sukchotrat, S.B. Kim, and F. Tsung. One-class classification-based control charts for multivariate process monitoring. *IIE transactions*, 42(2):107–120, 2009.
- [44] S.B. Kim, W. Jitpitaklert, and T. Sukchotrat. One-class classification-based control charts for monitoring autocorrelated multivariate processes. *Communications in Statistics—Simulation and Computation*(\mathbb{R}), 39(3):461–474, 2010.
- [45] W. Gani and M. Limam. Performance evaluation of one-class classification-based control charts through an industrial application. *Quality and Reliability Engineering International*, 29(6):841–854, 2013.
- [46] W. Gani and M. Limam. A one-class classification-based control chart using the-means data description algorithm. *Journal of Quality and Reliability Engineering*, 2014, 2014.
- [47] E.M. Maboudou-Tchao, I.R. Silva, and N. Diawara. Monitoring the mean vector with mahalanobis kernels. *Quality Technology & Quantitative Management*, 15(4):459–474, 2018.
- [48] J. Zhang, Z. Li, B. Chen, and Z. Wang. A new exponentially weighted moving average control chart for monitoring the coefficient of variation. *Computers & Industrial Engineering*, 78:205–212, 2014.
- [49] F.K. Wang, B. Bizuneh, and X.B. Cheng. One-sided control chart based on support vector machines with differential evolution algorithm. *Quality and Reliability Engineering International*, 35(6):1634–1645, 2019.
- [50] S. He, W. Jiang, and H. Deng. A distance-based control chart for monitoring multivariate processes using support vector machines. *Annals of Operations Research*, 263(1):191–207, 2018.
- [51] E.M. Maboudou-Tchao. Change detection using least squares one-class classification control chart. *Quality Technology & Quantitative Management*, 17(5):609–626, 2020.
- [52] M. Salehi, A. Bahreininejad, and I. Nakhai. On-line analysis of out-of-control signals in multivariate manufacturing processes using a hybrid learning-based model. *Neurocomputing*, 74(12):2083–2095, 2011. ISSN 0925-2312.

- [53] S. Hu and L. Zhao. A support vector machine based multi-kernel method for change point estimation on control chart. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pages 492–496, 2015.
- [54] W. Gani, H. Taleb, and M. Limam. Support vector regression based residual control charts. *Journal of Applied Statistics*, 37(2):309–324, 2010.
- [55] T. Sukchotrat, S.B. Kim, and F. Tsung. One-class classification-based control charts for multivariate process monitoring. *IIE Transactions*, 42(2):107–120, 2009.
- [56] D. Kakde, S. Peredriy, and A. Chaudhuri. A non-parametric control chart for high frequency multivariate data. In *2017 Annual Reliability and Maintainability Symposium (RAMS)*, pages 1–6. IEEE, 2017.
- [57] S. Jang, S.H. Park, and J.G. Baek. Real-time contrasts control chart using random forests with weighted voting. *Expert Systems with Applications*, 71:358–369, 2017. ISSN 0957-4174.
- [58] B.K. Issam and L. Mohamed. Support vector regression based residual mcusum control chart for autocorrelated process. *Applied Mathematics and Computation*, 201(1):565–574, 2008. ISSN 0096-3003.
- [59] S. Du, D. Huang, and J. Lv. Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines. *Computers & Industrial Engineering*, 66(4):683–695, 2013. ISSN 0360-8352.
- [60] J. Silva, O.B.P. Lezama, N. Varela, and M.S. Otero J.G. Guilianny, and E.S. Sanabria, and V.A. Rojas. U-control chart based differential evolution clustering for determining the number of cluster in k-means. In *International Conference on Green, Pervasive, and Cloud Computing*, pages 31–41. Springer, 2019.
- [61] C. Thirumalai, G. V. SaiSharan, K. V. Krishna, and K. J. Senapathi. Prediction of diabetes disease using control chart and cost optimization-based decision. In *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, pages 996–999, 2017.
- [62] G. Stefatos and A.B. Hamza. Statistical process control using kernel pca. In *2007 Mediterranean Conference on Control & Automation*, pages 1–6. IEEE, 2007.
- [63] P. Phaladiganon, S.B. Kim, V.C.P. Chen, and W. Jiang. Principal component analysis-based control charts for multivariate nonnor-

- mal distributions. *Expert Systems with Applications*, 40(8):3044–3054, 2013. ISSN 0957-4174.
- [64] J. Kullaa. Damage detection of the z24 bridge using control charts. *Mechanical Systems and Signal Processing*, 17(1):163–170, 2003. ISSN 0888-3270.
- [65] J.M. Lee, C.K. Yoo, S.W. Choi, P.A. Vanrolleghem, and I.B. Lee. Non-linear process monitoring using kernel principal component analysis. *Chemical Engineering Science*, 59(1):223–234, 2004. ISSN 0009-2509.
- [66] M. Ahsan, and H. Khusna M. Mashuri, M.H. Hidayatul, and Lee. Multivariate control chart based on kernel pca for monitoring mixed variable and attribute quality characteristics. *Symmetry*, 12(11):1838, 2020.
- [67] M. Ahsan, and D.D. Prastyo M. Mashuri, and H. Kuswanto, and H. Khusna. Multivariate control chart based on pca mix for variable and attribute quality characteristics. *Production & Manufacturing Research*, 6(1):364–384, 2018.
- [68] M. Mashuri, M. Ahsan, and D.D. Prastyo H. Kuswanto, and H. Khusna. Comparing the performance of t^2 chart based on pca mix, kernel pca mix, and mixed kernel pca for network anomaly detection. In *Journal of Physics: Conference Series*, volume 1752, page 012008. IOP Publishing, 2021.
- [69] W.J. Lee, and M.J. Triebe G.P. Mendis, and J.W. Sutherland. Monitoring of a machining process using kernel principal component analysis and kernel density estimation. *Journal of Intelligent Manufacturing*, 31(5):1175–1189, 2020.
- [70] J. Arkat, S.T.A. Niaki, and B. Abbasi. Artificial neural networks in applying mcusum residuals charts for ar(1) processes. *Applied Mathematics and Computation*, 189(2):1889–1901, 2007. ISSN 0096-3003.
- [71] S. Lee, M. Kwak, K.L. Tsui, and S.B. Kim. Process monitoring using variational autoencoder for high-dimensional nonlinear processes. *Engineering Applications of Artificial Intelligence*, 83:13–27, 2019.
- [72] S. Chen and J. Yu. Deep recurrent neural network-based residual control chart for autocorrelated processes. *Quality and Reliability Engineering International*, 35(8):2687–2708, 2019.
- [73] S.T.A. Niaki and B. Abbasi. Fault diagnosis in multivariate control charts using artificial neural networks. *Quality and Reliability Engineering International*, 21(8):825–840, 2005.

- [74] P. Chen, , Y. Li, K. Wang, M.J. Zuo, P.S. Heyns, and S. Baggerohr. A threshold self-setting condition monitoring scheme for wind turbine generator bearings based on deep convolutional generative adversarial networks. *Measurement*, 167:108234, 2021. ISSN 0263-2241.
- [75] G.A. Pugh. Synthetic neural networks for process control. *Computers & Industrial Engineering*, 17(1):24–26, 1989. ISSN 0360-8352.
- [76] T.F. Li, S. Hu, Z.Y. Wei, and Z.Q. Liao. A framework for diagnosing the out-of-control signals in multivariate process using optimized support vector machines. *Mathematical Problems in Engineering*, 2013, 2013.
- [77] R.S. Guh. Real-time recognition of control chart patterns in autocorrelated processes using a learning vector quantization network-based approach. *International Journal of Production Research*, 46(14):3959–3991, 2008.
- [78] M. Zaman and A. Hassan. Fuzzy heuristics and decision tree for classification of statistical feature-based control chart patterns. *Symmetry*, 13(1), 2021. ISSN 2073-8994.
- [79] W. Hachicha and A. Ghorbel. A survey of control-chart pattern-recognition literature (1991-2010) based on a new conceptual classification scheme. *Computers & Industrial Engineering*, 63(1):204–222, 2012. ISSN 0360-8352.
- [80] D.T. Pham and E. Oztemel. Control chart pattern recognition using combinations of multi-layer perceptrons and learning-vector-quantization neural networks. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 207(2):113–118, 1993.
- [81] C.S. Cheng. A neural network approach for the analysis of control chart patterns. *International Journal of Production Research*, 35(3):667–697, 1997.
- [82] A. Addeh, A. Khormali, and N. A. Golilarz. Control chart pattern recognition using rbf neural network with new training algorithm and practical features. *ISA transactions*, 79:202–216, 2018.
- [83] J. Yu, X. Zheng, and S. Wang. A deep autoencoder feature learning method for process pattern recognition. *Journal of Process Control*, 79:1–15, 2019.
- [84] J. Xu, H. Lv, Z. Zhuang, Z. Lu, D. Zou, and W. Qin. Control chart pattern recognition method based on improved one-dimensional con-

- volutional neural network. *IFAC-PapersOnLine*, 52(13):1537–1542, 2019.
- [85] W.A. Yang and W. Zhou. Autoregressive coefficient-invariant control chart pattern recognition in autocorrelated manufacturing processes using neural network ensemble. *Journal of Intelligent Manufacturing*, 26, 2015. ISSN 1161-1180.
- [86] D. Fuqua and T. Razzaghi. A cost-sensitive convolution neural network learning for control chart pattern recognition. *Expert Systems with Applications*, 150:113275, 2020.
- [87] D.T. Pham and M.A. Wani. Feature-based control chart pattern recognition. *International Journal of Production Research*, 35(7):1875–1890, 1997.
- [88] V. Ranaee, A. Ebrahimzadeh, and R. Ghaderi. Application of the psosvm model for recognition of control chart patterns. *ISA Transactions*, 49(4):577–586, 2010. ISSN 0019-0578.
- [89] C.J. Lu, Y.E. Shao, and P.H. Li. Mixture control chart patterns recognition using independent component analysis and support vector machine. *Neurocomputing*, 74(11):1908–1914, 2011. ISSN 0925-2312. Adaptive Incremental Learning in Neural Networks Learning Algorithm and Mathematic Modelling Selected papers from the International Conference on Neural Information Processing 2009 (ICONIP 2009).
- [90] S. Y. Lin, R. S. Guh, and Y. R. Shiue. Effective recognition of control chart patterns in autocorrelated data using a support vector machine based approach. *Computers & Industrial Engineering*, 61(4):1123–1134, 2011.
- [91] P. Xanthopoulos and T. Razzaghi. A weighted support vector machine method for control chart pattern recognition. *Computers & Industrial Engineering*, 70:134–149, 2014. ISSN 0360-8352.
- [92] X. Wang. Hybrid abnormal patterns recognition of control chart using support vector machining. In *2008 International Conference on Computational Intelligence and Security*, volume 2, pages 238–241, 2008.
- [93] V. Ranaee and A. Ebrahimzadeh. Control chart pattern recognition using a novel hybrid intelligent method. *Applied Soft Computing*, 11(2):2676–2686, 2011. ISSN 1568-4946. The Impact of Soft Computing for the Progress of Artificial Intelligence.

- [94] X. Zhou, P. Jiang, and X. Wang. Recognition of control chart patterns using fuzzy svm with a hybrid kernel function. *Journal of Intelligent Manufacturing*, 29(1):51–67, 2018.
- [95] H. De la Torre Gutierrez and D.T. Pham. Estimation and generation of training patterns for control chart pattern recognition. *Computers & Industrial Engineering*, 95:72–82, 2016. ISSN 0360-8352.
- [96] L. H. Chen and T. Y. Wang. Artificial neural networks to classify mean shifts from multivariate χ^2 chart signals. *Computers & Industrial Engineering*, 47(2-3):195–205, 2004.
- [97] C. S. Cheng and H. P. Cheng. Identifying the source of variance shifts in the multivariate process using neural networks and support vector machines. *Expert Systems with Applications*, 35(1-2):198–206, 2008.
- [98] R. S. Guh and Y.R. Shiue. An effective application of decision tree learning for on-line detection of mean shifts in multivariate control charts. *Computers & Industrial Engineering*, 55(2):475–493, 2008.
- [99] J. Yu, L. Xi, and X. Zhou. Identifying source (s) of out-of-control signals in multivariate manufacturing processes using selective neural network ensemble. *Engineering Applications of Artificial Intelligence*, 22(1):141–152, 2009.
- [100] E. Alfaro, J.L. Alfaro, M. Gamez, and N. Garcia. A boosting approach for understanding out-of-control signals in multivariate control charts. *International Journal of Production Research*, 47(24):6821–6834, 2009.
- [101] S. Verron, J. Li, and T. Tiplica. Fault detection and isolation of faults in a multivariate process with bayesian network. *Journal of Process Control*, 20(8):902–911, 2010.
- [102] S. G. He, Z. He, and G. A. Wang. Online monitoring and fault identification of mean shifts in bivariate processes using decision tree learning techniques. *Journal of Intelligent Manufacturing*, 24(1):25–34, 2013.
- [103] M. Carletti, C. Masiero, A. Beghi, and G. A. Susto. Explainable machine learning in industry 4.0: Evaluating feature importance in anomaly detection to enable root cause analysis. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 21–26. IEEE, 2019.
- [104] H. Song, Q. Xu, H. Yang, and J. Fang. Interpreting out-of-control signals using instance-based bayesian classifier in multivariate statistical process control. *Communications in Statistics-Simulation and Computation*, 46(1):53–77, 2017.

- [105] M. Salehi, A. Bahreininejad, and I. Nakhai. On-line analysis of out-of-control signals in multivariate manufacturing processes using a hybrid learning-based model. *Neurocomputing*, 74(12-13):2083–2095, 2011.
- [106] C. Zhao, H. Sun, and F. Tian. Total variable decomposition based on sparse cointegration analysis for distributed monitoring of nonstationary industrial processes. *IEEE Transactions on Control Systems Technology*, 28(4):1542–1549, 2019.
- [107] Q. Chen, U. Kruger, and A.Y.T. Leung. Cointegration testing method for monitoring nonstationary processes. *Industrial & Engineering Chemistry Research*, 48(7):3533–3543, 2009.
- [108] B.D. Ketelaere, K. Mertens, F. Mathijs, D.S. Diaz, and J.D. Baerde-maeker. Nonstationarity in statistical process control—issues, cases, ideas. *Applied Stochastic Models in Business and Industry*, 27(4):367–376, 2011.
- [109] J. Liu and D.S. Chen. Nonstationary fault detection and diagnosis for multimode processes. *AIChE journal*, 56(1):207–219, 2010.
- [110] T. Lazariv and W. Schmid. Surveillance of non-stationary processes. *AStA Advances in Statistical Analysis*, 103(3):305–331, 2019.
- [111] T. Lazariv and W. Schmid. Challenges in monitoring non-stationary time series. In *Frontiers in Statistical Quality Control 12*, pages 257–275. Springer, 2018.
- [112] P. Qiu. Big data? Statistical process control can help! *The American Statistician*, 74(4):329–344, 2020.
- [113] E. Tuv and G. Runger. Learning patterns through artificial contrasts with application to process control. *WIT Transactions on Information and Communication Technologies*, 29, 2003.
- [114] M.S. Reis and G. Gins. Industrial process monitoring in the big data/industry 4.0 era: From detection, to diagnosis, to prognosis. *Processes*, 5(3):35, 2017.
- [115] G. Capizzi and G. Masarotto. A least angle regression control chart for multidimensional data. *Technometrics*, 53(3):285–296, 2011.
- [116] F.M. Megahed, W.H. Woodall, and J.A. Camelio. A review and perspective on control charting with image data. *Journal of Quality Technology*, 43(2):83–98, 2011.
- [117] L. Zuo, Z. He, and M. Zhang. An ewma and region growing based control chart for monitoring image data. *Quality Technology & Quantitative Management*, 17(4):470–485, 2020.

- [118] H.D. Maragah and W.H. Woodall. The effect of autocorrelation on the retrospective x-chart. *Journal of Statistical Computation and Simulation*, 40(1-2):29–42, 1992.
- [119] J. Arkat, S.T.A. Niaki, and B. Abbasi. Artificial neural networks in applying mcusum residuals charts for ar (1) processes. *Applied Mathematics and Computation*, 189(2):1889–1901, 2007.
- [120] S.B. Kim, W. Jitpitaklert, S.K. Park, and S.J. Hwang. Data mining model-based control charts for multivariate and autocorrelated processes. *Expert Systems with Applications*, 39(2):2073–2081, 2012.
- [121] S. Cuentas, R. Peñabaena-Niebles, and E. Garcia. Support vector machine in statistical process monitoring: a methodological and analytical review. *The International Journal of Advanced Manufacturing Technology*, 91(1):485–500, 2017.
- [122] R.B. Chinnam and V.S. Kumar. Using support vector machines for recognizing shifts in correlated manufacturing processes. In *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No. 01CH37222)*, volume 3, pages 2276–2280. IEEE, 2001.
- [123] C.C. Hsu, M.C. Chen, and L.S. Chen. Integrating independent component analysis and support vector machine for multivariate process monitoring. *Computers & Industrial Engineering*, 59(1):145–156, 2010.
- [124] C.C. Hsu, M.C. Chen, and L.S. Chen. Intelligent ica-svm fault detector for non-gaussian multivariate process monitoring. *Expert Systems with Applications*, 37(4):3264–3273, 2010.
- [125] K.P. Tran, H.D. Nguyen, and S. Thomassey. Anomaly detection using long short term memory networks and its applications in supply chain management. *IFAC-PapersOnLine*, 52(13):2408–2412, 2019.
- [126] H.D. Nguyen, K.P. Tran, S. Thomassey, and M. Hamad. Forecasting and anomaly detection approaches using lstm and lstm autoencoder techniques with the applications in supply chain management. *International Journal of Information Management*, 57:102282, 2021.
- [127] C. Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, 2019.
- [128] K. Wang and W. Jiang. High-dimensional process monitoring and fault isolation via variable selection. *Journal of Quality Technology*, 41(3):247–258, 2009.

- [129] Y. Jin, S. Huang, G. Wang, and H. Deng. Diagnostic monitoring of high-dimensional networked systems via a lasso-bn formulation. *IIEE Transactions*, 49(9):874–884, 2017.
- [130] P. Qiu. Statistical process control charts as a tool for analyzing big data. In *Big and complex data analysis*, pages 123–138. Springer, 2017.
- [131] R. Sparks and S. Chakraborti. Detecting changes in location using distribution-free control charts with big data. *Quality and Reliability Engineering International*, 33(8):2577–2595, 2017.
- [132] F.M. Megahed and L.A. Jones-Farmer. Statistical perspectives on “big data”. In *Frontiers in statistical quality control 11*, pages 29–47. Springer, 2015.
- [133] P. Malaca, L.F. Rocha, D. Gomes, J. Silva, and G. Veiga. Online inspection system based on machine learning techniques: real case study of fabric textures classification for the automotive industry. *Journal of Intelligent Manufacturing*, 30(1):351–361, 2019.
- [134] W.H. Woodall and D.C. Montgomery. Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, 46(1):78–94, 2014.
- [135] E. Bochinski, T. Senst, and T. Sikora. Hyper-parameter optimization for convolutional neural network committees based on evolutionary algorithms. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3924–3928. IEEE, 2017.
- [136] H. Trittenbach, K. Böhm, and I. Assent. Active learning of svdd hyperparameter values. In *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 109–117. IEEE, 2020.
- [137] V.V. Trinh, K.P. Tran, and T.T. Huong. Data driven hyperparameter optimization of one-class support vector machines for anomaly detection in wireless sensor networks. In *2017 International Conference on Advanced Technologies for Communications (ATC)*, pages 6–10. IEEE, 2017.
- [138] J. Wu, S. P. Chen, and X.Y. Liu. Efficient hyperparameter optimization through model-based reinforcement learning. *Neurocomputing*, 409:381–393, 2020.
- [139] S. Hosseini and B. Mohammad Hasani Zade. New hybrid method for attack detection using combination of evolutionary algorithms, svm, and ann. *Computer Networks*, 173:107168, 2020.

- [140] I. Žliobaitė, M. Pechenizkiy, and J. Gama. An overview of concept drift applications. *Big data analysis: new algorithms for a new society*, pages 91–114, 2016.
- [141] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia. A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, 46(4):1–37, 2014.
- [142] G. Shmueli and S.E. Fienberg. Current and potential statistical methods for monitoring multiple data streams for biosurveillance. In *Statistical methods in counterterrorism*, pages 109–140. Springer, 2006.
- [143] F. Castanedo. A review of data fusion techniques. *The scientific world journal*, 2013, 2013.
- [144] M. Zhang, Y. Yuan, R. Wang, and W. Cheng. Recognition of mixture control chart patterns based on fusion feature reduction and fireworks algorithm-optimized msvm. *Pattern Analysis and Applications*, 23(1): 15–26, 2020.
- [145] M. Zhang, X. Zhang, H. Wang, G. Xiong, and W. Cheng. Features fusion exacton and kelm with modified grey wolf optimizer for mixture control chart patterns recognition. *IEEE Access*, 8:42469–42480, 2020.
- [146] Y. Umeda, J. Kaneko, and H. Kikuchi. Topological data analysis and its application to time-series data analysis. *Fujitsu Scientific & Technical Journal*, 55(2):65–71, 2019.
- [147] B.M. Colosimo and M. Pacella. A comparison study of control charts for statistical monitoring of functional data. *International Journal of Production Research*, 48(6):1575–1601, 2010.
- [148] J. Liu, J. Chen, and D. Wang. Wavelet functional principal component analysis for batch process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 196:103897, 2020.
- [149] M. Flores, R. Fernández-Casal S. Naya, S. Zaragoza, P. Raña, and J. Tarrío-Saavedra. Constructing a control chart using functional data. *Mathematics*, 8(1):58, 2020.
- [150] G. Yu, C. Zou, and Z. Wang. Outlier detection in functional observations with applications to profile monitoring. *Technometrics*, 54(3): 308–318, 2012.
- [151] Z. He, L. Zuo, M. Zhang, and F.M. Megahed. An image-based multivariate generalized likelihood ratio control chart for detecting and diagnosing multiple faults in manufactured products. *International Journal of Production Research*, 54(6):1771–1784, 2016.

- [152] K. He, and L. Zuo M. Zhang, T. Alhwiti, and F.M. Megahed. Enhancing the monitoring of 3d scanned manufactured parts through projections and spatiotemporal control charts. *Journal of Intelligent Manufacturing*, 28(4):899–911, 2017.
- [153] S.E. Stankus and K.K. Castillo-Villar. An improved multivariate generalised likelihood ratio control chart for the monitoring of point clouds from 3d laser scanners. *International Journal of Production Research*, 57(8):2344–2355, 2019.
- [154] Y. Okhrin, W. Schmid, and I. Semeniuk. Monitoring image processes: Overview and comparison study. In *International Workshop on Intelligent Statistical Quality Control*, pages 143–163. Springer, 2019.
- [155] Y. Okhrin, W. Schmid, and I. Semeniuk. New approaches for monitoring image data. *IEEE Transactions on Image Processing*, 30:921–933, 2020.
- [156] Y. Yuan and L. Lin. Self-Supervised Pre-Training of Transformers for Satellite Image Time Series Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020.
- [157] Phuong Hanh Tran, Cédric Heuchenne, and Sébastien Thomassey. An anomaly detection approach based on the combination of lstm autoencoder and isolation forest for multivariate time series data. In *Proceedings of the 14th International FLINS Conference on Robotics and Artificial Intelligence (FLINS 2020)*, pages 18–21. World Scientific, 2020.
- [158] S.J. Sheather and J.S. Marron. Kernel quantile estimators. *Journal of the American Statistical Association*, 85(410):410–416, 1990.
- [159] H. Qiu, J. Lee, J. Lin, and Gang G. Yu. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of sound and vibration*, 289(4-5):1066–1090, 2006.