

Received March 8, 2021, accepted March 28, 2021, date of publication April 2, 2021, date of current version April 13, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3070659

A Multi-Step Anomaly Detection Strategy Based on Robust Distances for the Steel Industry

KISAN SARDA¹, (Student Member, IEEE),
ANTONIO ACERNESE¹, (Student Member, IEEE), **VITTORIO NOLÈ²**,
LEONARDO MANFREDI², **LUCA GRECO³**, **LUIGI GLIELMO¹**, (Senior Member, IEEE),
AND CARMEN DEL VECCHIO¹, (Member, IEEE)

¹Department of Engineering, University of Sannio, 82100 Benevento, Italy.

²Siderpotenza—Pittini Group, 85100 Potenza, Italy.

³Department of Law, Economics, Management and Quantitative Methods (DEMM), University of Sannio, 82100 Benevento, Italy.

Corresponding author: Kisan Sarda (kirasarda@unisannio.it)

This work was supported by the Programma Operativo Nazionale (PON) Project “Piattaforma decisionale Manutenzione Processo e produttivo (PROMPT)” through the Italian Ministry for University and Research under Grant CUP F26C18000180005.

ABSTRACT Steel making industries exhibit extreme working conditions characterized by high temperature, pressure, and production speed as well as intense throughput. Due to high economic and energy investments of the overall production process, an intense and expensive preventive maintenance program is adopted to avoid breakdowns. Steel making process would greatly benefit from a predictive maintenance module able to detect incoming faults from data process analysis. However, due to intense preventive maintenance, available data recording process operations enclose only a few samples of fault events, avoiding the efficient application of classical data driven anomaly detection models. In an attempt to overcome the above mentioned limits, we report the outcome of an industrial research project on data-driven anomaly detection in a steel making production process. The study assesses a fault detection strategy for rotating machines in the hot rolling mill line: we developed an automatic two-step strategy, which combines two statistical methods over the available data set: more precisely, the combination of Re-weighted Minimum Covariance Determinant estimator and Hidden Markov Models helped identify working conditions in a drive reducer of a hot steel rolling mill line and automatically isolate signs of decreasing performance or upcoming failures. The proposed strategy has been validated on real data collected in a steel making plant placed in the South of Italy.

INDEX TERMS Anomaly detection, rolling mills, reweighted minimum covariance determinant (RMCD), Hidden Markov Model (HMM), robust statistics.

I. INTRODUCTION

The steel making process forges a collection of raw materials, including steel scrap, carbon, and limestone, into steel bars with different diameter sizes. In order to obtain the highest production rates, today's steel industries handle heavier loads and faster velocity than ever before. Indeed, continuous production processes working in these harsh environments characterized by high temperatures, pressures, and production speeds need to flow uninterrupted [1]: unexpected equipment failures are expensive and potentially catastrophic, resulting in severe risk for plant operators, unplanned

The associate editor coordinating the review of this manuscript and approving it for publication was Baoping Cai.

production downtime, costly replacement of parts. To reduce breakdown impacts on production, steel making companies adopt preventive maintenance strategies substituting equipment long before the end of their useful life [2], [3]. Apart from scheduled maintenance, production line reliability may be increased adopting fully automatic monitoring and fault diagnosis systems. Indeed, a major task in steel making industries is the detection of faults, whereas neither an analytical description of faults and process models nor the collection of typical breakdown patterns exists [4].

Fault detection methods can be roughly classified into two broad categories: model-based methods and data-based methods (also referred to as history-based). The first relies on a mathematical model of the system under analysis; faults are

detected by computing the error between the estimated model and the actual values of the process variable. However, this method requires a detailed theoretical model the derivation of which is a complex and time-consuming task, in particular when production lines are equipped with rotating machines. Data-based methods have been largely adopted since they do not require any specific a priori knowledge of the system characteristics; they only rely on the processing of available data measurements and statistical techniques [5], [6]. A fundamental requirement for the effective implementation of data-based method is the availability of relevant, information rich training data. However, in the particular environment of high risk/high intensity industries -such as steel industries- the availability of a large training data set is in fact impracticable since machine faults are prevented and unlikely to occur. Lack of breakdown samples induces to address further investigation towards the development of efficient methods to detect anomalies in production data, rather than to predict fault events [4].

In an attempt to settle an anomaly detection module in the steel production line, we developed a two-step scheme, relying on complementary data-driven techniques. Specifically, a preliminary anomaly detection phase has been carried out based on Minimum Covariance Determinant (MCD), a simple robust multivariate outlier technique. Secondly, a more complex method is proposed based on Hidden Markov Models (HMM), targeting a more precise identification of the anomalies. Both techniques essentially rely on the fusion of information received from sensors to obtain one reliable measure and a corresponding cut-off level; new data coming from the operating plant are compared with the cut-off level to be classified as anomalies, normal operating conditions, or intermediate operating stages.

The preliminary anomaly detection method of our scheme uses multivariate analyses techniques which exploit the nature of anomalies, i.e. measures aside from main features of the majority of the data, exhibiting their own pattern or even no pattern at all. These measures are also referred to as outliers; in what follows we will refer to anomalies and outliers interchangeably. Various methods for detecting multivariate outliers have been proposed in the literature (see [7], [8] for an overview); out of these, Principal Component Analysis (PCA) and Independent Component Analysis (ICA) have been largely adopted in industrial anomaly detection with some applications to steel making industry [9]–[12]. While intuitive and popular, PCA has been developed under the assumption that the extracted principal components (PCs) have a Gaussian distribution; in practice, PCA is not often applicable, and it has been proved to provide unreliable monitoring results [8]. ICA-based monitoring has been developed to overcome PCA limits and it has been successfully adopted to analyze non-Gaussian data in the steel casting process [8]. However, ICA also showed several drawbacks: computational load, unavailable criteria for sorting independent components, and solutions changing at each computation since ICA is initialized randomly.

Among various available robust techniques for estimation of multivariate location and scatter and outliers detection, we focus on the MCD estimator [13]–[15]. Briefly, the idea behind the MCD is to discard a fraction of most distant data points; the data to be retained are chosen as those whose sample variance-covariance matrix has the lowest determinant. These data are expected to be the closest to each other and not contaminated by outliers. In order to increase efficiency, a re-weighting step is usually performed: observations with squared distances from the MCD fit above a given threshold are trimmed. Finally, anomalies conditions are identified looking for multivariate outliers by the inspection of univariate distances. In particular, data points whose distance is larger than a fixed threshold can be pointed as potential outliers.

Multivariate analysis provides binary information (i.e. normal/anomalous) on production process operating conditions. It is a simple yet useful instrument to build up a fully automated process data classification. However, a complex production process such as rolling mills shows a transition from normality to failure through degradation of process performances over time. Indeed, some changes in process variables (such as vibration, oil temperature, motor electrical absorption) may alter the operating conditions without immediately leading to a breakdown; however, if unreported, the persistent shift over time to a range of intermediate efficiency states will conduct to a failure. Hence, to monitor and predict several latent states of the activity of the production process along time we fit an Hidden Markov Model (HMM) over distances. The HMM provides a dynamic classification of distances and in particular of large distances corresponding to outliers into the latent states. Furthermore, the HMM returns the probability of moving towards degradation stages to the final inferred state denoting failure. HMM represent a well-known statistical tool to monitor different stages of the process and for automatic outliers detection [16]–[18]. Here we mainly rely on HMM to deal with the continuous-in-time nature of the available measurements. Furthermore, we also stress that effective outliers detection rules should be driven from suitable robust methods well suited to handle outliers [19].

We applied the two-step scheme described above to a steel making industry placed in the south of Italy. We focused on anomaly detection in rolling mills of the metal sheet forming processes. This problem has been rather studied in the literature. Machine learning techniques have been more successfully applied to faults affecting product qualities [20]–[23]; when applied to anomaly detection methods, most of the proposed strategies rely on classical machine learning methodologies applied to rotating machines (as in [24], [25]) hence they are strongly limited by the absence of data on breakdown events in the steel making process. To overcome this limit faults are usually simulated, for example, corrupting a vibration measurement with an additive train of pulses [25]. A residual based analysis has been adopted in [4]; authors rely on model identification of main rolling mills components and classify anomalies accordingly. Despite some intriguing

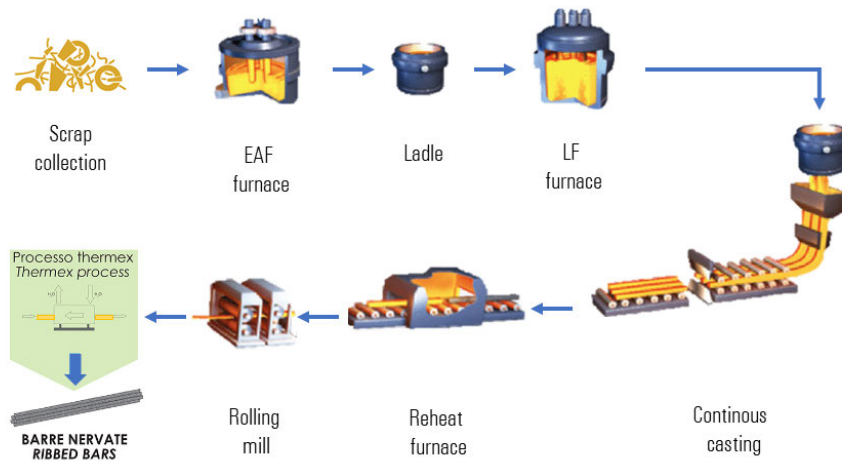


FIGURE 1. Steel making process.

results, this methodology suffers from oversimplification of plant activity and it is impracticable in a long chain of rolling mills. Conversely, the two-stage scheme we developed relies on robust measures; in fact, the HMM is identified over the distances of the data points from the robust fit stemming obtained through RMCD estimation, allowing a dynamic strategy that takes into account the time-varying nature of the data and infers several latent states in the production process. Robust distance fuses the information received from sensors displaced on the production process, hence obtaining one reliable measure and the corresponding trade-off, easily implementable and adaptable to all process stages. The procedure has been validated through production data from the working plant.

Among our main contributions we consider:

- The identification of a metric fusing sensors' measures in one reliable feature.
- The design of an automatic procedure for anomaly detection in large data sets with strongly unbalanced clusters.
- The extension of the static anomaly detection method to an HMM-based method providing additional information on the evolution of anomaly severity in time.
- The validation of the proposed methodologies to rolling mills of a steel production plant, using data recorded from Pittini, a steel making plant placed in the South of Italy.

II. DEVICE DESCRIPTION AND PROBLEM SETTING

A. STEEL MAKING PROCESS

Steel making is a multistage process including steel making, continuous casting, and rolling stages (see Figure 1). The process starts with the melting of raw materials into the electric arc furnace (EAF) at 1600°C . The molten steel is continuously supplied into the casting lines, in which the liquid steel is “guided” through rolls; the output of this stage is a semi-finished product, namely the billet.

Billets are sequentially rolled in multiple rolling mills, i.e., a chain of rotating machines that progressively stretches the bars and reduces their diameters. The mill drive train supplies mechanical energy to the top and bottom driven rollers of the rolling mill stand; it is configured with an electrical motor, a gear drive, and a pinion stand, all connected by couplings; the combination of these elements is also referred to as a cage (Figure 2).

A steel making plant generally produces several types of bars differing for their diameters (also denoted as profiles); each profile is worked by a subset of the available milling machines. For instance, Pittini plant produces 25 profiles (ϕ , in mm) each one worked by a specific combination of the 18 cages comprising the milling line.

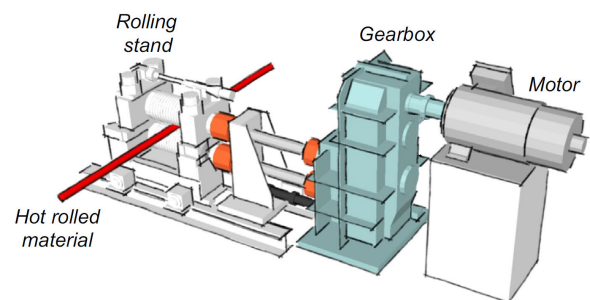


FIGURE 2. 3D rendering example of the reversible rolling mill to be studied, Figure 1 in [25].

In order to obtain the highest production rates, today's rolling mills handle heavier loads and faster velocities than ever before, leading to high temperatures, pressures, and degradation. The layout of a rolling mill is highly vulnerable to sudden breakdowns: if a fault occurs in one of the operating cages, it may cause injuries to operators, damages to other mechanical equipment, and relevant economical losses. For these reasons, faults are usually prevented by scheduling

frequent and costly maintenance interventions which usually comprise the replacement of critical components much earlier than the end of their useful life.

B. AVAILABLE SIGNALS AND MONITORING SYSTEM

The core system to be studied consists of an electric motor and a gear reducer, represented in Figure 3, serving the blooming mill of the cage. The following are cage parameters monitored in Pittini's plant; they represent as well a list of the most frequently recorded data rolling mills:

- Three vibration signals measured by three accelerometers placed on the shell of the drive reducer, sampled at 10 kHz.
- Oil pressure in the drive reducer, acquired at 1 Hz.
- Two temperature signals measured by two probes placed in the opposite sides of the reducer, sampled at 1 Hz.
- Current absorbed by the motor driving the mill, with the sampling frequency of 10 Hz.
- Revolution speed of the motor, measured in Round Per Minute (RPM) and acquired at 10 Hz.
- Torque of the motor sampled at 10 Hz.

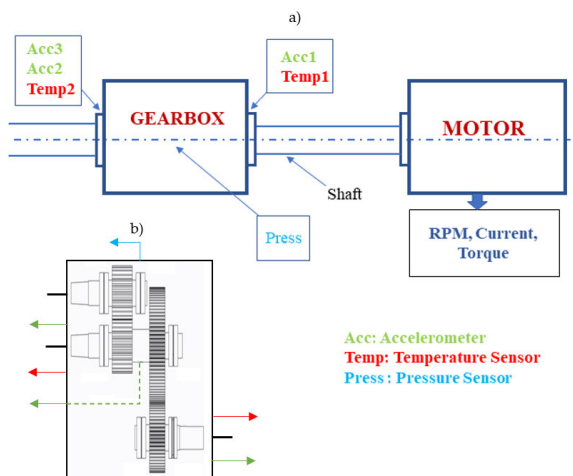


FIGURE 3. Electric motor and gear scheme with the corresponding set of available sensors and their positions (a). Gearbox structure, with the positions of the sensors (b).

The set of the above mentioned sensors is replicated on 6 of the 18 cages of the plant, i.e., the monitoring system of the rolling stand supervises the activity of the 6 most used rolling machines. The recording system is completely automatic: it acquires and stores on a server the data corresponding to one registration every 10 produced billets, i.e., the data are stored only if the plant is producing. In particular, starting from 5 sec when the billet is loaded to all cages involved in processing the ongoing profile, the acquisition lasts for 15 sec and outputs a data set with a number of rows equal to the number of samples and a number of columns equal to the total number of sensors. The corresponding impact in terms of memory load is about 100 MB/billet.

C. STRATEGY FOR ANOMALY DETECTION

In this section, we describe our two steps methodology for anomaly detection in rolling mills, schematically is represented in Figure 4; the theoretical background supporting each step is detailed in Section III while examples of its application to Pittini's plant are in Section IV.

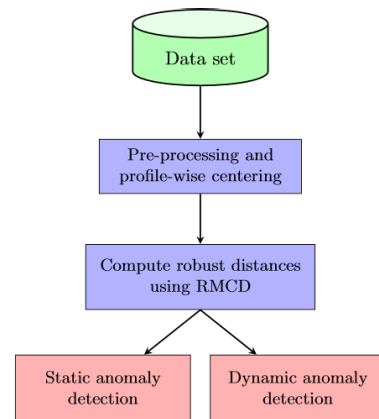


FIGURE 4. Flow-chart of the proposed methods.

The first preparatory step is the computation of **residuals** on the acquired measures. This step is peculiar to Pittini's data set, nevertheless, it may apply to several production sights. The supervisory system records process parameters of all worked profiles; according to its production plan, each profile is milled in some cages at peculiar reference values of the production parameters. Indeed, different profiles are worked at different nominal values of temperature, pressure, and line velocity. The first objective of our work is to obtain one reference parameter -fusing information received from sensors-independent from billets profiles. To this aim a conditional mean has been evaluated: for each profile and for each feature the mean of all recorded data is computed. In order words, a **profile-wise centering operation** has been performed in order to avoid the effects due to data acquisitions from distinct profiles. With a slight abuse of terminology, we refer to the record of the profile-wise centered data as the data set of residuals; this will be input to subsequent phases. Since the data are supposed to be contaminated by the occurrence of several outliers, we perform a centering operation based on a robust mean, rather than on the classical sample mean. Here, we use the Huber estimate of location [19]. The use of the sample mean is deprecated since it would be attracted by outliers unavoidably and all residuals severely shifted. As a consequence, a large rate of genuine observation could be transformed into anomalous.

The second step concerns the **robust fitting** of multivariate location and scatter over the p -dimensional residuals evaluated using the RMCD estimator. Next, relying on the robust estimate, a **static outliers detection test** is performed for each data point at an overall confidence level $\alpha = 0.01$. The testing procedure is based on distances from the robust fit. Finally, an HMM is fitted to the robust distances in order to

take into account time dependence among subsequent measurements (i.e. a **dynamic outliers detection** is performed). One of the inferred HMM states is expected to collect outliers and denote a failure state.

The analysis of robust distances through the RMCD leads to appropriate testing rules that are meant to detect inconsistent data. Although efficient, this kind of strategy is basic since it only deals with the pair of latent conditions anomaly/not anomaly and it is likely to suffer a large number of false alarms (false positives). On the contrary, the transition from a normal operating condition towards faults can be identified fitting an HMM over such distances evaluated on the entire data set. Moreover, the outliers detection rule from RMCD is static: it does not take into account the sequential nature of the data, whereas the analysis on robust distances made by fitting an HMM is dynamic, the evolution of the clustering process being governed by a first-order Markov chain.

III. MULTIVARIATE STATISTICAL METHODS

A. STATIC ANOMALY DETECTION

Notation: We denote by \mathbb{R} and \mathbb{N}_+ the sets of real and positive natural numbers, respectively. Uppercase bold variables \mathbf{X} represent matrices, lowercase bold ones \mathbf{x} represent vectors, that are supposed to be column, i.e., $\mathbf{x} = (\cdot, \dots, \cdot)^\top$ is a column vector.

Let $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)^\top$ be a sample of size n from a p -variate Normal distribution $\mathcal{Y} \sim \mathcal{N}_p(\mu, \Sigma)$, where μ is the sample mean, Σ is the sample covariance matrix, and $p \in \mathbb{N}_+$ is the number of features, i.e., $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{ip}) \in \mathbb{R}^p$, $\mu \in \mathbb{R}^p$, and $\Sigma \in \mathbb{R}^{p \times p}$, respectively.

Definition 1 (Mahalanobis distance): The Mahalanobis distance (MD) between the i th sample $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{ip})$ and the sample mean $\mu \in \mathbb{R}^p$, with respect to the covariance matrix $\Sigma \in \mathbb{R}^{p \times p}$, is

$$d_i := d(\mathbf{y}_i; \mu, \Sigma) = \sqrt{(\mathbf{y}_i - \mu)^\top \Sigma^{-1} (\mathbf{y}_i - \mu)}. \quad (1)$$

A traditional approach to detect outliers from a data set is based on the MD evaluated on the sample mean $\hat{\mu}$ and the sample covariance matrix $\hat{\Sigma}$ of the observed data \mathbf{Y} . In particular, a sample \mathbf{y} is considered anomalous if $d(\mathbf{y}; \hat{\mu}, \hat{\Sigma})$ is larger than some fixed threshold. Under the assumed p -variate Normal model with known μ and Σ , the squared MD obeys a chi-squared distribution with p degrees of freedom, i.e., $d^2(\mathbf{Y}; \mu, \Sigma) \sim \chi_p^2$. A frequent choice [19] consists in using as a threshold the 0.975- or 0.99-level quantile of the χ_p^2 distribution. However, it is well known that traditional approaches in computing $\hat{\mu}$ and $\hat{\Sigma}$ are not robust to the outliers, indeed they are biased by outliers themselves.

This motivates the exploration of methods to robustly estimate the parameters of the p -variate Normal distribution. The MCD is one of the most popular methods. The method is very popular as it relies on good asymptotic properties [26] and the availability of efficient algorithms for its computation [14]. The idea behind MCD is to discard a fraction ϵ of the most

distant data points [19], with ϵ denoting the trimming level determining the size of the considered subsample. The data to be retained are chosen as the set of size $\bar{n} := \lfloor n(1 - \epsilon) \rfloor$ whose covariance matrix has the lowest determinant (where $\lfloor \cdot \rfloor$ indicates the floor function), i.e., the data that are most close to each other are less likely to be outlying.

Given data \mathbf{Y}_i , the MCD estimator is

$$\begin{aligned} \hat{\mu} &= \frac{1}{\sum_i z_i} \sum_i z_i \mathbf{y}_i, \\ \hat{\Sigma} &= \frac{c(p, \epsilon)}{\sum_i z_i - 1} \sum_i z_i (\mathbf{y}_i - \hat{\mu})(\mathbf{y}_i - \hat{\mu})^\top, \end{aligned} \quad (2)$$

where \mathbf{z} is a binary vector whose components z_i equal to zero in case of a trimmed observation, and the consistency factor $c(p, \epsilon)$ is used to make the MCD consistent at the p -Normal model, by inflating the covariance matrix based on the selected subset [19].

The MCD can be evaluated according to an iterative algorithm that assures a monotonic decrease of the determinant of the estimated covariance matrices, i.e., at each iteration $j = 1, 2, \dots$, $|\hat{\Sigma}^j| \leq |\hat{\Sigma}^{j-1}|$, where $|\cdot|$ here is the determinant, and the superscript is used to represent the iteration. Therefore, the sequence of determinants is expected to converge in a finite number of iterations. The complete MCD procedure is summarized in Algorithm 1.

Algorithm 1 MCD algorithm

Require: $\hat{\mu}_0, \hat{\Sigma}_0$ be initial estimate based on a subsample of size \bar{n}, ϵ, n, p
Ensure: $\hat{\mu}_{MCD}, \hat{\Sigma}_{MCD}$
1: $\hat{\mu}_{MCD} \leftarrow \hat{\mu}_0, \hat{\Sigma}_{MCD} \leftarrow \hat{\Sigma}_0$
2: **repeat**
3: **for** $i = 1, 2, \dots, n$ **do**
4: $d_i = \sqrt{(\mathbf{y}_i - \hat{\mu}_{MCD})^\top \hat{\Sigma}_{MCD}^{-1} (\mathbf{y}_i - \hat{\mu}_{MCD})}$
5: **end for**
6: Sort distances $d_{(1)} \leq d_{(2)} \dots \leq d_{(n)}$ in non increasing order and take a subsample $\bar{\mathbf{Y}}$ of size \bar{n} based on the lowest distances
7: Estimate $\hat{\mu}_{MCD}$ and $\hat{\Sigma}_{MCD}$ using (2)
8: **until** convergence

In order to increase the efficiency of the method in the finite horizon, a reweighting step is usually performed, leading to the reweighted MCD (RMCD) method. The reweighting step is performed downstream of Algorithm 1 and works as follows. Given the distances $d_{i,MCD} = \sqrt{(\mathbf{y}_i - \hat{\mu}_{MCD})^\top \hat{\Sigma}_{MCD}^{-1} (\mathbf{y}_i - \hat{\mu}_{MCD})}$, $i = 1, 2, \dots, n$ at convergence, computed as described in Algorithm 1, then the observations \mathbf{y}_i for which $d_{i,MCD}^2 > \chi_{p,q}^2$ are trimmed, where $\chi_{p,q}^2$ is the q -quantile of the χ_p^2 distribution. Let $\bar{\epsilon} \leq \epsilon$ be the rate of observation trimmed in the reweighting step. Then,

the RMCD is obtained as the sample covariance matrix over the non-trimmed set of data inflated by $c(\bar{\epsilon}, p)$.

Let $(\hat{\mu}_{RMCD}, \hat{\Sigma}_{RMCD})$ denote the RMCD estimate. Then, an observation \mathbf{y} is labeled as anomalous if

$$d^2(\mathbf{y}_i; \hat{\mu}_{RMCD}, \hat{\Sigma}_{RMCD}) > \chi_{p;1-\alpha}^2. \quad (3)$$

It is worth noting that the problem of mislabeled instances can be further faced with additional well-known methods. For example, when the number of instances is relatively large, to contain the number of false positives (FPs), we suggest controlling multiplicity by some correction procedure, such as the False Discovery Rate (FDR) [27], as we show in the experiments section.

B. DYNAMIC ANOMALY DETECTION

The anomaly detection strategy described so far can be considered as a static analysis as it identifies only two states in the process, i.e., no failure and failure, thus overlooking any kind of time dependencies. However, real conditions of manufacturing processes exhibit continuous trends over time, starting from healthy states and dropping toward some failure conditions. In light of this dependency of any machinery health's conditions on the time, it would be desirable to implement a dynamic analysis on the data. To this aim, we here provide a general method based on HMM that outputs a set of latent states along with an estimate of their transition probability distribution.

An HMM is a framework for representing probability distributions over sequences of observations. It involves two interconnected models: i) the state model, consisting of a discrete-time, discrete-state, first-order Markov chain, with states $z_t \in \{1, 2, \dots, k\}$ and transition probability $P(z_t|z_{t-1})$, and ii) the observation model, whose underlying dynamics is given by $P(y_t|z_t)$, where y_t is an observation at time t . Thus, in this model, an observation y_t is produced by a stochastic process \mathcal{Z} , but the state z_t of this process cannot be directly observed, i.e., it is hidden [28]. The stochastic process \mathcal{Z} is assumed to satisfy the Markov property, i.e., given z_{t-1} , the current hidden state z_t is independent of all the states prior to $t - 1$. The corresponding joint distribution of a sequence of states $z_{1:\tau}$ and observations $y_{1:\tau}$ can be factored as

$$\begin{aligned} P(z_{1:\tau}, y_{1:\tau}) &= P(z_1)P(y_1|z_1) \prod_{t=2}^{\tau} P(z_t|z_{t-1})P(y_t|z_t) \\ &= P(z_1) \prod_{t=2}^{\tau} P(z_t|z_{t-1}) \prod_{t=1}^{\tau} P(y_t|z_t). \end{aligned}$$

To summarize, an HMM can be defined by the tuple $(k, n, \mathbf{P}, \mathbf{Q}, \mathbf{P}_0)$, where k is the number of possible hidden states, assumed to be discrete, and n is the number of distinct observations. Moreover, \mathbf{P} and \mathbf{Q} determine the underlying transition models for the states and the observations, respectively, with $\mathbf{P}_{i,j} = P(z_t = i|z_{t-1} = j)$ and $\mathbf{Q}_{i,j} = P(y_t = i|z_t = j)$. The last element, \mathbf{P}_0 , represents the initial state distribution.

The author in [29] introduced the idea that HMMs should be characterized by three fundamental problems:

- 1) (Likelihood) Given an HMM $\lambda = (\mathbf{P}, \mathbf{Q})$, and an observation sequence $\mathbf{Y} = (y_1, y_2, \dots, y_n)^\top$, determine the likelihood $P(\mathbf{Y}|\lambda)$.
- 2) (Decoding) Given an observation sequence \mathbf{Y} and an HMM $\lambda = (\mathbf{P}, \mathbf{Q})$, discover the best hidden state sequence $\mathbf{z} = (z_1, z_2, \dots, z_n)^\top$.
- 3) (Learning) Given an observation sequence \mathbf{Y} , the set of states k in the HMM and the initial state distribution \mathbf{P}_0 , learn the HMM parameters \mathbf{P} and \mathbf{Q} .

Problem 1 aims at computing the likelihood of a particular observation sequence \mathbf{Y} . Because the hidden state sequence corresponding to \mathbf{Y} is not known, one possibility is to sum over all possible hidden state sequences, i.e., $P(\mathbf{Y}|\lambda) = \sum_{\mathbf{z}} P(\mathbf{Y}|\mathbf{z}, \lambda)P(\mathbf{z}|\lambda)$. This procedure is extremely computationally inefficient because there are k^n possible hidden sequences, thus requiring an exponential time in the number of states. A more efficient solution is called forward algorithm and it is a kind of dynamic programming algorithm that computes and stores in a table the partial probabilities for each state-time pair, and iteratively computes the following probability:

$$\alpha_t(j) = \sum_{i=1}^k \alpha_{t-1}(i)P_{i,j}b_j(\mathbf{y}_t), \quad (4)$$

where $\alpha_t(j) = P(\mathbf{Y}, z_t = j|\lambda)$, and $b_j(\mathbf{y}_t)$ is the state observation likelihood of the observation y_t given the current state j . After having initialized the probability $\alpha_1(j) = P_{0,j}b_j(\mathbf{y}_1)$, $1 \leq j \leq k$, the forward step (4) can be recursively applied for all states and observations in order to compute the marginal $P(\mathbf{Y}|\lambda) = \sum_{i=1}^k \alpha_n(i)$.

An extended version of the forward algorithm can be used for the learning problem. In particular, if we design a backward procedure that computes the backward probability of seeing the observations from time $t + 1$ to the end, i.e., $\beta_t(i) := P(y_{t+1}, \dots, y_n|z_t = i, \lambda)$, we can solve the problem of training both \mathbf{P} and \mathbf{Q} . This algorithm is known as the Baum-Welch algorithm [30], and it is a special case of the well-known Expectation-Maximization (EM) algorithm [31]. The recursive backward rule is given by

$$\beta_t(i) = \sum_{j=1}^k \beta_{t+1}(j)P_{i,j}b_j(\mathbf{y}_{t+1}). \quad (5)$$

The most probable state sequence \mathbf{z} for an observed sequence \mathbf{Y} can be obtained exhaustively by deriving all the possible state sequences, generating $P(\mathbf{z}|\mathbf{Y}, \lambda)$, and finally obtaining the \mathbf{z} for which this probability is maximum. Even if this approach may be feasible for relatively small \mathbf{Y} , it becomes computationally expensive as n increases. To efficiently obtain the most probable state path for a given observation sequence, i.e., to solve the decoding problem, we can use the Viterbi algorithm [32] that is a dynamic programming method that efficiently computes the best path probability

by substituting the sum operation (over the previous path probabilities) with a maximization. It is worth noting that, for page limit reasons, we do not formally state the three algorithms as they are standard methods used to solve the HMM problems, and we remind to [28], [30], [32] for their formal definitions.

Based on the above discussions, we can now cast the HMM problems to our dynamic anomaly detection task. In particular, assuming that the hidden states represent the health conditions of the monitored system, it would be important to estimate the probability to transition from one state to another, given a sequence of observations, i.e., problem 3. Moreover, given an estimate of the model and a new observation sequence, it would be desirable to assign to such sequence a corresponding hidden state sequence, e.g., using the Viterbi algorithm.

In summary, the proposed methodology performs as follows:

- 1) compute the robust distances $d(\mathbf{y}_i; \hat{\mu}_{RMCD}, \hat{\Sigma}_{RMCD})$, $i = 1, \dots, n$;
- 2) fit an HMM model using the Baum-Welch algorithm;
- 3) given a new observation, compute its robust distance from the RMCD estimate $d(\mathbf{y}; \hat{\mu}_{RMCD}, \hat{\Sigma}_{RMCD})$ and predict the corresponding hidden state using the Viterbi algorithm.

In all the aforementioned methods, the number k of hidden states is assumed to be known. In real applications it is usually determined according to information criteria, such as the Akaike or the Bayesian information criteria (AIC and BIC) [33].

IV. RESULTS AND DISCUSSIONS

In this Section, we present and discuss the results obtained applying the proposed methodologies on real data collected from the steel making production process described in Section I. In particular, in the following, we show the results obtained from data stored at 1 Hz and at 10 kHz; we do not show the results obtained with 10 Hz data since they did not exhibit any anomalous patterns during the observation time. Moreover, during the observation period lasting three months, we observed only one significant downtime caused by a fault on the 18th cage; the overall number of billets produced during the observation time is equal to 5634. This unique occurrence is a rare event, taking into account the frequent preventive maintenance interventions scheduled to maximize the reliability of the production process. In light of the above considerations, we use the HMM-based approach to incorporate time dependencies into anomaly detection model and provide an automatic methodology able to estimate an health model based on real data. The application of the proposed methodology on production data shows benefits derived from our approach.

The data set adopted in the following analysis has been pre-processed to remove NaN data and anomalous zero values; moreover, as discussed in Section II-C, a robust

profile-wise centering step has been performed to obtain a data set independent on the produced profiles.

A. LOW FREQUENCY DATA

In this Subsection, we show the results of the static and dynamic anomaly detection methods applied to the data acquired at 1 Hz, namely two temperatures and one pressure transducer. Figure 5 displays the robust distances along with the detected anomalies for an overall level $\alpha = 0.01$. The points above the cut-off line are those detected as outliers. In particular, there are 360 flagged points (outliers) of which 108 with profile ϕ_8 .

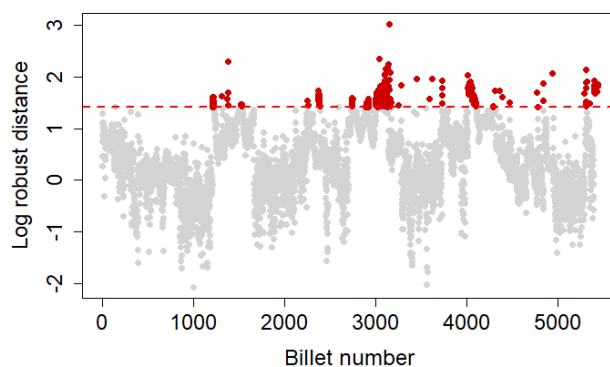


FIGURE 5. Anomaly detection using RMCD+FDR on 1 Hz data. The fault corresponds to the samples around billet 3000 and is correctly identified as anomaly.

According to the steps highlighted in Section II-C, we first obtained robust residuals by inspection of the different profiles. Then, we applied the RMCD with an initial trimming level of 25%. The reweighting step leads to the deletion of about 17% of the data. Trimmed observations may give an indication of the presence of outliers and the actual level of contamination in the data but the outliers detection process is supposed to be performed separately, based on the current robust fit and formal rules as in (3). Figure 5 displays the robust distances along with the detected anomalies for an overall level $\alpha = 0.01$. The points above the cut-off line are those detected as outliers.

As far as the dynamic anomaly detection analysis is concerned, we designed an HMM for a Gamma observations distribution and log link, as the chi-squared distribution belongs to the Gamma family, and the logarithmic link allows to estimate a positive mean. We chose the number of states according to the BIC criterion, and we considered a multi-start approach based on fifty random starts in order to make the final maximum likelihood estimate independent from the initialization. As we can see in Table 1, the BIC criterion suggests a model with $k = 6$ hidden states.

The results from the HMM analysis over the distances from the RMCD fit are given in Figure 6 and Figure 7, which display the distribution of the robust distances by the six inferred states and their classification on a log scale, respectively. The group with the largest modal distance (whose

TABLE 1. Selection of the number of latent states using the BIC applied to 1 Hz data.

k	BIC	k	BIC
2	33288.90	5	23124.00
3	26.962.87	6	22765.14
4	24.346.98	7	23074.13

TABLE 2. Transition matrix for 1 Hz data.

	to z_1	to z_2	to z_3	to z_4	to z_5	to z_6
from z_1	0.884	0.029	0.026	0.001	0.059	0.001
from z_2	0.111	0.790	0.026	0.004	0.061	0.007
from z_3	0.006	0.008	0.939	0.020	0.024	0.003
from z_4	0.003	0.003	0.020	0.973	0.000	0.001
from z_5	0.033	0.005	0.078	0.002	0.880	0.002
from z_6	0.000	0.635	0.051	0.000	0.000	0.314

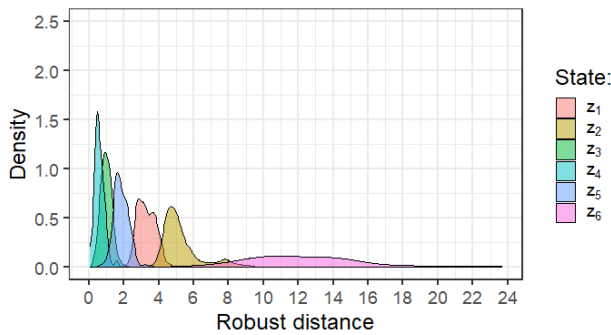


FIGURE 6. 1 Hz data. Distribution of robust distances by latent states.

label is 6) should be interpreted as the fault alert state. It is composed of 18 points, 16 of which are of $\phi 8$. On the contrary, the groups with the lowest modal distance can be seen as a normal condition state, whereas the other represents intermediate situations that characterize the transitions from normality to anomaly. Clearly, the HMM analysis of the robust distances led to a more plausible classification and dynamic description of the evolution of the process over time. Table 2 gives the fitted transition probabilities. In particular, we notice that the probabilities to have a fault when the process is in any of the other states are rather small, with a total probability of 0.014. Furthermore, the probability that the process from the fault alert condition of state 6 returns in the previous pre-alert state labeled 2 is 0.635. This finding may suggest the potential occurrence of false alerts. There is also a remarkable persistence in the fault alarm state, as the probability to remain in state 6 is 0.314.

We also ran a classical multivariate Gaussian HMM to the original data, still taking into account profile-wise centering. In this case, the BIC criterion leads to choose $k = 8$, which is a larger number of states underlying a more complex to interpret data structure (not only from a strict visualizing point of view). From the inspection of conditional means and variance-covariance matrices, it turns out that the failure state, characterized by the more extreme values for the second temperature sensor and the pressure sensor, is composed

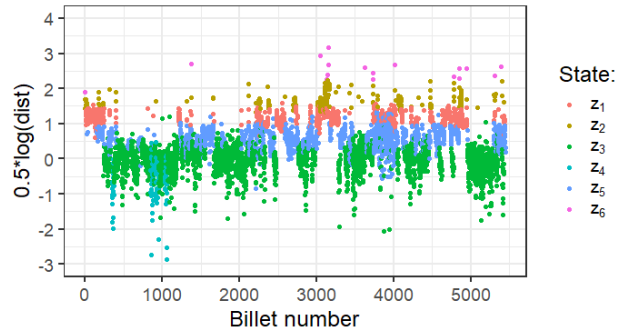


FIGURE 7. 1 Hz data. Robust distances by latent states on log scale.

TABLE 3. Transition matrix for 10 kHz data.

	to z_1	to z_2	to z_3	to z_4	to z_5
from z_1	0.739	0.223	0.000	0.037	0.000
from z_2	0.445	0.540	0.010	0.001	0.004
from z_3	0.017	0.000	0.795	0.188	0.000
from z_4	0.000	0.018	0.164	0.819	0.000
from z_5	0.070	0.014	0.000	0.000	0.916

of 400 points. In addition, there are at least a couple of alert states. We can state that the HMM evaluated over robust distances returns a more feasible description of the data structure and effective detection of the failure state. On the contrary, the classical HMM suffers from masking effects, since it is not able to detect outliers properly with a lower number of states and even with $k = 8$ is likely to flag many false positives.

B. HIGH FREQUENCY DATA

The same analysis has been performed on the data generated by the three accelerometers. The data generated for each billet have been summarized with the mean values, that are immediately affected by anomalies in the production process. Figure 8 shows the distribution of robust distances among profiles; the horizontal line gives the cut-off to detect outliers. It is evident that the majority of outliers belong to profiles of $\phi 8$: 158 over a total of 381 detected anomalies. In particular, the more extreme points in the first box-plot are those corresponding to the observed fault.

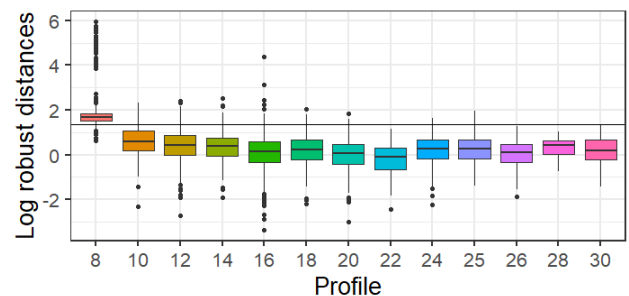


FIGURE 8. 10 kHz data. Distribution of robust distances over profiles.

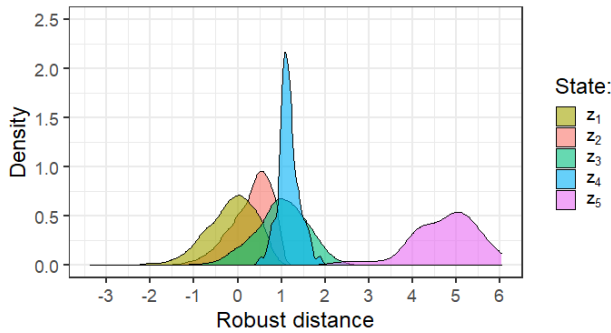


FIGURE 9. 10 kHz data. Distribution of robust distances by latent states.

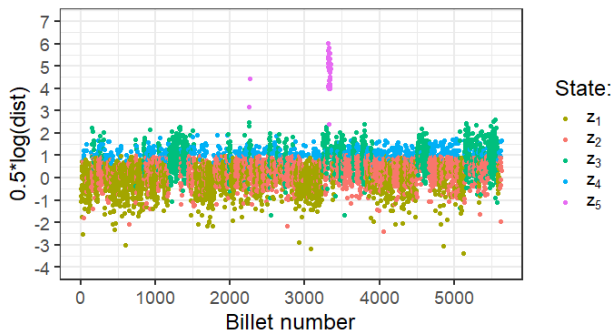


FIGURE 10. 10 kHz data. Robust distances by latent states on log scale.

By running the HMM, the BIC criterion identifies five latent states. In this example, the HMM has been fitted after that the squared robust distances have been transformed according to the Wilson-Hilferty formula [34], which allows to transform χ_p^2 distributed variate towards normality. Therefore, a Gaussian HMM has been fitted because the Gamma HMM showed severe convergence issues. The BIC criterion selected a model with five latent states. The distribution of distances within each state is displayed in Figure 9, whereas their classification is given in Figure 10. We notice that the fifth cluster includes the most incipient fault samples: 34 over 36 anomalies concern $\phi 8$ profile. The output from HMM clearly reduced the risk of false alarms with respect to the former static analysis. The entries in Table 3 give the transition probabilities: the probability of remaining in the fault state is 91.6%.

In order to better assess the performance of the proposed technique, we also fitted a Gaussian multivariate HMM to the high frequency profile-wise centered values. The classical procedure returned a similar structure with the same number of latent states and a similar classification into states. The main difference concerns the fitted conditional probability to remain in the fault state that now is lower than its robust counterpart and equal to 0.876. Here, the classical procedure is able to deal with outliers since they are very far from the bulk of the data and characterized by huge values. Nevertheless, the underestimation of the conditional probability to remain

in the alarm state could badly affect the prediction of future anomalies.

C. COMBINING DATA ANALYSIS: 1 Hz AND 10 kHz DATA

We here consider the combination of the above analyzed data, i.e., temperature, pressure, and vibration signals obtained through robust distances. Let us focus on the results stemming from the HMM analysis over the robust distances. The BIC criterion suggested to select seven hidden states. The seven clusters inferred from the fitted HMM are given in Figure 11. It is quite evident that the state labeled 1 includes very large outliers. In this latent state 39 points have been classified of which 37 are from profile $\phi 8$. The classification is displayed in Figure 12. These results are in strong agreement with those stemming from the previous analyses. In this example, fitting the classical HMM model under the assumption of multivariate normality of the data at hand does not lead to select a feasible number of latent states, with $k \geq 10$. It is evident that such a result makes any interpretation effort challenging and likely to be not useful. This means that the classical multivariate HMM is not able to cope with outliers, if not at the cost of an extremely complex underlying structure that can undo any prediction effort.

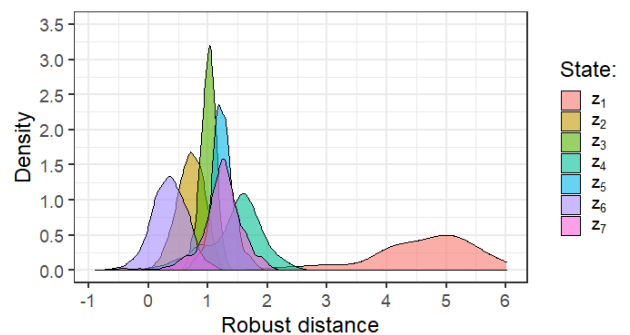


FIGURE 11. Combined data. Distribution of robust distances by latent states.

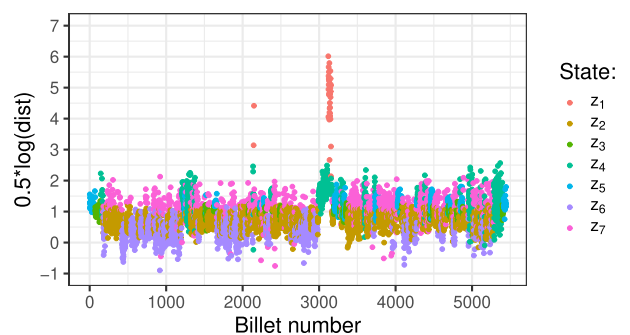


FIGURE 12. Combined data. Robust distances by latent states on log scale.

D. PREDICTION AND ALARM DETECTION

In order to investigate and assess the predictive capability of the proposed strategy, we decided to split the data into

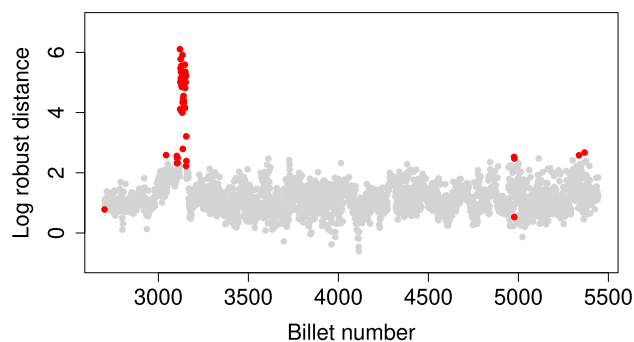


FIGURE 13. Detected anomalies in the test data (red circles).

two groups: the first 2700 observations are used as train data, whereas the remaining 2744 as test data. The static outliers detection based on the reweighted MCD leads to flag about 20% of the test data as outliers. On the contrary, the HMM analysis detects only 49 points, after classifying robust distances into seven latent states: the points are displayed in Figure 13. The method is able to detect the occurred anomalies, despite few false alarms.

V. CONCLUSION

In this work we have proposed a two-step scheme, relying on two different methodologies (static and dynamic), for fault detection for rolling mill in a steel production plant, using real field data. A preliminary fault detection phase has been carried out based on robust MDs and on the RMCD. This method appears to be very effective for the detection of faulty samples, and it results computationally efficient. Secondly, based on the robust distances obtained in the previous phase, a more complex method is then used to confirm the detection of the fault and provide additional information about the probability of transitioning among the latent states, i.e., the HMM-based method can potentially predict the evolution over time of the status of the monitored system. In the results section, we proved the effectiveness of the proposed procedures when applied to real data.

We believe that the proposed methodology is attractive because it is straightforwardly implementable yet sufficiently accurate. Future work will cover the extension of the proposed methods to the frequency domain. Moreover, it will be investigated the application of the proposed anomaly detection method to an automatic plant's supervisor systems such as a digital twin.

REFERENCES

- [1] C. Neef, S. Hirzel, and M. Arens, "Industry 4.0 in the European iron and steel industry: Towards an overview of implementations and perspectives," Fraunhofer Inst. Syst. Innov. Res. ISI, Karlsruhe, Germany, 2018.
- [2] N. Naidu, "Development of mathematical models for optimal preventive maintenance policy in a steel industry: Six sigma approach," *Int. J. Qual. Res.*, vol. 5, no. 3, pp. 1–10, 2011.
- [3] N. Sakib and T. Wuest, "Challenges and opportunities of condition-based predictive maintenance: A review," *Procedia CIRP*, vol. 78, pp. 267–272, Jan. 2018.

- [4] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger, and H. Efendic, "Residual-based fault detection using soft computing techniques for condition monitoring at rolling mills," *Inf. Sci.*, vol. 259, pp. 304–320, Feb. 2014.
- [5] M. G. Pecht and M. Kang, "Machine learning: Anomaly detection," in *Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things*. Piscataway, NJ, USA: IEEE, 2019, pp. 131–162.
- [6] S. Jeschke, C. Brecher, T. Meisen, D. Özdemir, and T. Eschert, "Industrial Internet of Things and cyber manufacturing systems," in *Industrial Internet of Things*. Cham, Switzerland: Springer, 2017, pp. 3–19.
- [7] M. Fahim and A. Sillitti, "Anomaly detection, analysis and prediction techniques in IoT environment: A systematic literature review," *IEEE Access*, vol. 7, pp. 81664–81681, 2019.
- [8] R. Chalapathy, A. K. Menon, and S. Chawla, "Robust, deep and inductive anomaly detection," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Cham, Switzerland: Springer, 2017, pp. 36–51.
- [9] H. Wang, M. J. Bah, and M. Hammad, "Progress in outlier detection techniques: A survey," *IEEE Access*, vol. 7, pp. 107964–108000, 2019.
- [10] Z. Ji, X. Zhang, and C. Wang, "Monitoring of continuous steel casting process based on independent component analysis," in *Proc. Chin. Control Decis. Conf.*, May 2010, pp. 3920–3923.
- [11] P. Zhou, R. Zhang, J. Xie, J. Liu, H. Wang, and T. Chai, "Data-driven monitoring and diagnosing of abnormal furnace conditions in blast furnace ironmaking: An integrated PCA-ICA method," *IEEE Trans. Ind. Electron.*, vol. 68, no. 1, pp. 622–631, Jan. 2021.
- [12] J.-M. Lee, C. Yoo, and I.-B. Lee, "Statistical process monitoring with independent component analysis," *J. Process Control*, vol. 14, no. 5, pp. 467–485, Aug. 2004.
- [13] P. J. Rousseeuw, "Multivariate estimation with high breakdown point," *Math. Statist. Appl.*, vol. 8, nos. 283–297, p. 37, 1985.
- [14] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, Aug. 1999.
- [15] P. J. Rousseeuw and M. Hubert, "Anomaly detection by robust statistics," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 8, no. 2, p. e1236, Mar. 2018.
- [16] D. Gao, M. K. Reiter, and D. Song, "Behavioral distance measurement using hidden Markov models," in *Proc. Int. Workshop Recent Adv. Intrusion Detection*. Berlin, Germany: Springer, 2006, pp. 19–40.
- [17] P. Smyth, "Hidden Markov models for fault detection in dynamic systems," *Pattern Recognit.*, vol. 27, no. 1, pp. 149–164, Jan. 1994.
- [18] J. Li, W. Pedrycz, and I. Jamal, "Multivariate time series anomaly detection: A framework of hidden Markov models," *Appl. Soft Comput.*, vol. 60, pp. 229–240, Nov. 2017.
- [19] A. Farcomeni and L. Greco, *Robust Methods for Data Reduction*. Boca Raton, FL, USA: CRC Press, 2015.
- [20] B. Song, H. Shi, S. Tan, and Y. Tao, "Multi-subspace orthogonal canonical correlation analysis for quality related plant wide process monitoring," *IEEE Trans. Ind. Informat.*, early access, Aug. 7, 2020, doi: 10.1109/TII.2020.3015034.
- [21] B. Song, H. Yan, H. Shi, and S. Tan, "Multisubspace elastic network for multimode quality-related process monitoring," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 5874–5883, Sep. 2020.
- [22] B. Song, X. Zhou, H. Shi, and Y. Tao, "Performance-indicator-oriented concurrent subspace process monitoring method," *IEEE Trans. Ind. Electron.*, vol. 66, no. 7, pp. 5535–5545, Jul. 2019.
- [23] C. Del Vecchio, G. Fenu, F. A. Pellegrino, M. Di Foggia, M. Quatralo, L. Benincasa, S. Iannuzzi, A. Acernese, P. Corrao, and L. Glielmo, "Support vector representation machine for superalloy investment casting optimization," *Appl. Math. Model.*, vol. 72, pp. 324–336, Aug. 2019.
- [24] H. Luo, K. Li, O. Kaynak, S. Yin, M. Huo, and H. Zhao, "A robust data-driven fault detection approach for rolling mills with unknown roll eccentricity," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 6, pp. 2641–2648, Nov. 2020.
- [25] M. Farina, E. Osto, A. Perizzato, L. Piroddi, and R. Scattolini, "Fault detection and isolation of bearings in a drive reducer of a hot steel rolling mill," *Control Eng. Pract.*, vol. 39, pp. 35–44, Jun. 2015.
- [26] R. W. Butler, P. L. Davies, and M. Jhun, "Asymptotics for the minimum covariance determinant estimator," *Ann. Statist.*, vol. 21, no. 3, pp. 1385–1400, Sep. 1993.
- [27] A. Cerioli and A. Farcomeni, "Error rates for multivariate outlier detection," *Comput. Statist. Data Anal.*, vol. 55, no. 1, pp. 544–553, Jan. 2011.

[28] W. Zucchini, I. L. MacDonald, and R. Langrock, *Hidden Markov Models for Time Series: An Introduction Using R*. Boca Raton, FL, USA: CRC Press, 2017.

[29] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

[30] L. E. Baum, "An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," in *Proc. 3rd Symp. Inequalities*, O. Shisha, Ed. New York, NY, USA: Academic, 1972, pp. 1–8.

[31] A. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. Roy. Stat. Soc. B, Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.

[32] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 260–269, Apr. 1967.

[33] N. Dridi and M. Hadzagic, "Akaike and Bayesian information criteria for hidden Markov models," *IEEE Signal Process. Lett.*, vol. 26, no. 2, pp. 302–306, Feb. 2019.

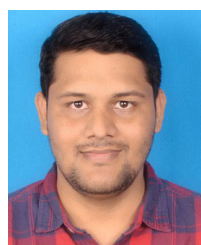
[34] E. B. Wilson and M. M. Hilferty, "The distribution of chi-square," *Proc. Nat. Acad. Sci. USA*, vol. 17, no. 12, p. 684, 1931.



LEONARDO MANFREDI received the degree in mechanical engineering from the University of Basilicata, in 2005. From 2005 to 2006, he worked as an Electromechanical Installer on behalf of Mazzetti Renato Srl. Since 2006, he has been working with Siderpotenza—Pittini Group Plant, as the Manager of the rolling mill.



LUCA GRECO received the degree from the University of Sannio, in 2001, and the Ph.D. degree in statistics from the Department of Statistical Sciences, University of Padova, in 2005. In 2006, he was an Associate Researcher with the Department of Statistics, Ca' Foscari University of Venice. Since December 2006, he is has been an Assistant Professor with the University of Sannio. His current research interests include likelihood theory, robust methods, weighted likelihood, flexible models, multivariate statistics, data reduction, and mixture models. He is currently a member of the Italian Statistical Society. He received the National Professional Qualification (ASN) to be an Associate Professor in statistics, in 2013 and 2019, and the National Professional Qualification (ASN) to be a Full Professor, in 2020.



KISAN SARDA (Student Member, IEEE) received the bachelor's degree in electronics and telecommunication engineering and the master's degree in electrical engineering with the focus on control systems from the University of Mumbai, India, in 2013 and 2017, respectively. He is currently pursuing the Ph.D. degree in information technologies. His current research interests include deep learning, system identification and control in the manufacturing field, prognostics and predictive maintenance, and Boolean networks.



ANTONIO ACERNESE (Student Member, IEEE) received the bachelor's and master's degrees in electronics engineering in automation and telecommunications from the University of Sannio, Benevento, Italy, in 2015 and 2017, respectively, with a focus on automation, where he is currently pursuing the Ph.D. degree in information technologies engineering. His current research interests include reinforcement learning, optimization and control in the manufacturing field, prognostics and predictive maintenance, prediction, and control methods.



VITTORIO NOLÈ worked as the Technical Director with Team Progetto Corsa, dealing with car development, data acquisition, and analysis, from 2008 to 2016. Since 2016, he has been working with the Siderpotenza—Pittini Group Plant, as a Mechanical Designer and moved to the Maintenance Engineering Department, in 2018.



LUIGI GLIELMO (Senior Member, IEEE) was born in 1960. He received the Laurea degree in electronic engineering and the Ph.D. degree in automatic control. He taught at the University of Palermo, the University of Naples Federico II, and the University of Sannio at Benevento, where he is currently a Professor of automatic control. He has coauthored more than 160 papers on international archival journals or proceedings of international conferences, co-edited two books, and holds three patents. His research interests include singular perturbation methods, model predictive control methods, automotive controls, deep brain stimulation modeling and control, smart-grid control, and systems biology. He was the General Co-Chair of the 2019 European Control Conference, Naples, Italy. He has been the Chair of the IEEE Control Systems Society Technical Committee on Automotive Controls. He seated in the editorial boards of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He also serves as an Associate Editor for the on-line journal IEEE CONTROL SYSTEMS LETTERS.



CARMEN DEL VECCHIO (Member, IEEE) received the Laurea degree in management engineering from the University of Naples Federico II, Italy, in 1999, and the Ph.D. degree in control and computer science from the University of Sannio, Italy, in 2004. Since 2011, she has been an Assistant Professor with the Engineering Department, University of Sannio. Her research interests include optimization and control of interconnected processes through various theoretical instruments, such as decisions Markov chains, probabilistic Boolean networks, machine, and reinforcement learning, and applications to several fields, including manufacturing, smart grids, medicine, and biology.

...