processes

Review

Data-Driven Process Monitoring and Fault Diagnosis: A Comprehensive Survey

Afrânio Melo [1,2,*], Maurício Melo Câmara [1,3] and José Carlos Pinto [1]

[1] Programa de Engenharia Química/COPPE, Universidade Federal do Rio de Janeiro, Cidade Universitária, Rio de Janeiro CEP 21941-972, Brazil; mauricio@peq.coppe.ufrj.br (M.M.C.); pinto@peq.coppe.ufrj.br (J.C.P.)
[2] Centro de Pesquisas Leopoldo Americo Miguez de Mello—CENPES, Petrobras—Petróleo Brasileiro SA, Rio de Janeiro CEP 21941-915, Brazil
[3] OPTIMATECH Ltda., Cidade Universitária, Rio de Janeiro CEP 21941-614, Brazil
* Correspondence: afranio.melo@petrobras.com.br

Abstract: This paper presents a comprehensive review of the historical development, the current state of the art, and prospects of data-driven approaches for industrial process monitoring. The subject covers a vast and diverse range of works, which are compiled and critically evaluated based on the different perspectives they provide. Data-driven modeling techniques are surveyed and categorized into two main groups: multivariate statistics and machine learning. Representative models, namely principal component analysis, partial least squares and artificial neural networks, are detailed in a didactic manner. Topics not typically covered by other reviews, such as process data exploration and treatment, software and benchmarks availability, and real-world industrial implementations, are thoroughly analyzed. Finally, future research perspectives are discussed, covering aspects related to system performance, the significance and usefulness of the approaches, and the development environment. This work aims to be a reference for practitioners and researchers navigating the extensive literature on data-driven industrial process monitoring.

Keywords: process systems engineering; process monitoring; fault detection and diagnosis; multivariate statistics; latent variable modeling; machine learning

check for updates

Citation: Melo, A.; Câmara, M.M.; Pinto, J.C. Data-Driven Process Monitoring and Fault Diagnosis: A Comprehensive Survey. *Processes* **2024**, *12*, 251. https://doi.org/10.3390/pr12020251

Academic Editors: Jie Zhang and Wei Sun

Received: 7 December 2023
Revised: 11 January 2024
Accepted: 17 January 2024
Published: 24 January 2024

1. Introduction

The ongoing advancements in society demand continuous improvements in safety, efficiency, and sustainability within process industries. To tackle these challenges, companies must have a comprehensive understanding of process behaviors, enabling them to quickly detect, identify and fix malfunctions. The rise in digitalization has provided access to vast amounts of data collected from various sensors, creating an opportunity for informed decision making. Research on data-driven process monitoring aims to develop methods that enable real-time use of process data in the most effective manner, with the goal of ensuring healthy performance.

Achieving this objective involves various aspects, such as mathematical modeling, data processing, benchmarking methodologies, software development, and practical implementation in industrial settings. The literature on data-driven process monitoring has seen significant growth in recent years, with numerous studies addressing these aspects. In particular, many review articles and overviews have been published, typically focusing on specific topics. For instance, Yu and Zhang [1] discuss the challenges of deep learning modeling; Kong [2] explores the application of latent variable models; Zhao [3] examines non-stationary process monitoring; Peres and Fogliatto [4] review strategies for variable selection; and Melo et al. [5] analyze open benchmarks for comparing methodologies.

Despite the wealth of existing works, to the best of our knowledge, the literature lacks a comprehensive survey encompassing at once the following gaps: (i) compiling and critically analyzing previous reviews related to the topic; (ii) presenting the most popular mathematical models in a didactic manner, ensuring the inclusion of most significant models; (iii) conducting a thorough examination of often overlooked aspects, such as practical implementation and software development; and (iv) covering the historical progression, the current state of the art, and future perspectives of the field.

The primary objective of the present paper is to provide a survey of the research field of process monitoring, specifically addressing the four aforementioned points. By addressing these points in a single document, this review seeks to offer a holistic understanding of the field and its fundamental aspects. The next sections of the article are organized as follows. Section 2 provides a broad overview of the historical development of process monitoring. Detailed discussions on mathematical modeling techniques are presented, specifically multivariate statistical models (Section 3) and machine learning approaches (Section 4). Section 5 focuses on the exploration, characterization, and treatment of process data. Section 6 reviews computational tools and openly available benchmarks in the literature. Section 7 addresses the practical implementation of process monitoring in industrial settings. Section 8 explores future perspectives and challenges in the field. Section 9 concludes the article. Additionally, Appendixes A and B include a compilation of relevant resources such as review papers, overviews, comparative studies, books, and commercial tools related to process monitoring in the field of process systems engineering (PSE). These resources can serve as references for further exploration of specific topics.

2. Historical Remarks

Process monitoring is a term widely used in both daily application and engineering practice, encompassing various aspects such as the supervision of production by operators. However, as industries have evolved, specific methodologies have been developed to systematize this procedure.

Shewhart [6,7], during his tenure at Bell Telephone Laboratories, introduced the concept of control charts. These graphical tools, rooted in statistical principles, enable the monitoring of significant process variables to detect events associated with special cause variations, distinct from common cause variations that are inherent and expected during process operation. The development of control charts is widely recognized as a pivotal moment in the establishment of statistical process monitoring [8].

Another significant breakthrough occurred with the introduction of the T^2 statistic by Hotelling [9]. This statistical tool allowed for the simultaneous analysis of multiple variables and their interactions, marking the inception of the field of multivariate statistical process monitoring [10]. This pioneering work enabled a comprehensive approach to process monitoring by considering the interconnectedness among various variables.

Despite Hotelling's work in 1947, the widespread adoption of multivariate monitoring applications occurred primarily in the late 1980s. This can be attributed to two key factors: (i) the increased accessibility and utilization of computers and (ii) the proposal of applying latent variable modeling, a particular class of multivariate statistical methodologies, to process monitoring problems [11]. These methodologies have become the prevailing approach from the 1990s onwards, marking a paradigm shift that has been extensively documented in numerous review and overview articles from that era [12–14].

Since the 2000s, the rapid advancement of computational capacity has led to the growing prominence of machine learning approaches for multivariate process monitoring. The ability of machine learning algorithms to process and analyze vast amounts of data has contributed to their increasing relevance in the field of multivariate monitoring, enabling the identification of complex relationships and patterns that were previously challenging to detect [15,16].

In recent years, there has been a growing tendency to classify multivariate statistical techniques as machine learning algorithms [16,17]. However, despite this trend, in the PSE literature, they are conventionally categorized separately [10,15,18,19]. This separation arises from the historical development of multivariate statistical models in fields such as statistics and chemometrics, along with the association of machine learning methodologies with the broader field of computer science. In this study, we present the two approaches in dedicated sections. Section 3 provides an overview of the main techniques developed within the framework of multivariate statistics, while Section 4 focuses on the techniques derived from the machine learning literature. By separating these discussions, we aim to offer a clear and comprehensive examination of each approach within the context of its historical development. However, it is important to keep in mind that nuances and gray areas may emerge: while multivariate techniques can be interpreted as learning algorithms, many machine learning methodologies inherently possess a statistical foundation.

3. Multivariate Statistics

In multivariate statistics, multiple variables and their interactions are analyzed simultaneously. In this context, latent variable modeling stands out as the most popular multivariate statistical approach for applications in process monitoring [16]. This approach involves transforming the raw data's original variables into a new set of variables that explicitly reveal the underlying, hidden (latent) characteristics present in the data. The choice of technique depends on the specific characteristics one wishes to uncover. Some commonly used techniques are principal component analysis (PCA), partial least squares (PLS) and canonical correlation analysis (CCA) [20–23]. These methods are widely applied in industrial process monitoring applications because they generate orthogonal latent variables that eliminate the linear correlations between the original variables, a common characteristic in industrial data [24].

Latent variable modeling has gained such historical significance that several works which aimed to present the literature on data-driven process monitoring solely focused on this methodology [14,18,20,25–28]. These works may suggest two implicit viewpoints: either the entire process monitoring field is exclusively constituted by these methods, or these are the only methods worthy of attention.

Only MacGregor and Cinar [26] provided a justification for this preference:

> "*In this paper, we focus on the proper use of data-driven (empirical) models for the monitoring, control and optimization of processes. In particular, we focus on latent variable models because, as we will show, they provide the proper structure to allow them to be built from plant data and be used for monitoring, control and optimization*" [26].

Throughout their article, the authors justified their choice by arguing that latent variable modeling methods are the only ones that can model both the predictor (**X**) and predicted (**Y**) portions of a dataset. According to the authors, other regression techniques would solely model the **Y** space, which would be equivalent to assuming that the **X** space has full rank, a hypothesis rarely met in real-world scenarios. As a consequence, latent variable methods would be the most suitable for handling missing data issues and the only data-driven monitoring methodology that could provide unique, interpretable, and causal models.

Several objections can be made to this viewpoint. The shift to the latent variable space is not the only way to address collinearity in the **X** matrix. The nonlinearity of some regression methods can itself mitigate or eliminate this issue, and even in the context of linear methods, other strategies exist, including subset selection and penalization methods such as ridge regression and Lasso [29]. Data reconciliation techniques can also be employed, as they allow for the treatment of fluctuations and uncertainties in both **X** and **Y** [30–33]. The aim of these objections is not to diminish the importance of latent variable models but rather to highlight that their prevalence in the process monitoring literature may be partially attributed to historical biases. Recognizing this fact can help foster openness to new opportunities and better understand the shift in discourse towards buzzwords such as machine learning and artificial intelligence, as discussed in Section 4.

In the remaining part of this section, a detailed account of the main techniques used in multivariate statistics is presented from a historical and didactic perspective. The theoretical descriptions emphasize the assumptions that each technique makes regarding the nature of the data. While these assumptions are essential in characterizing the use of the techniques, they are often overlooked by practitioners.

3.1. PCA

3.1.1. Fundamentals

Principal component analysis (PCA) was originally introduced by Pearson [34]; however, the application in the field of process monitoring was proposed much later by Wise et al. [11]. PCA is a technique that produces a set of orthogonal linear combinations of the original variables, with the goal of selecting a subset of these combinations that effectively captures and summarizes the variability of the data. The first assumption of the technique refers to the approach used for deriving the latent variables:

PCA, Assumption 1. *The latent variables can be obtained by performing a linear transformation on the dataset*.

Let $\mathbf{X} \in \mathbb{R}^{n \times m}$ be a dataset, where each row corresponds to an observation and each column corresponds to a measured variable. The linear transformation performed by PCA can be represented as a change in basis:

$$\mathbf{T} = \mathbf{X}\mathbf{P}, \tag{1}$$

where the columns of $\mathbf{P} \in \mathbb{R}^{m \times m}$, denoted by $\{\mathbf{p_1}, \ldots, \mathbf{p_m}\}$, are referred to as principal components. These components represent a new basis for expressing each observation (i.e., row) of the matrix $\mathbf{X}$. On the other hand, the columns of $\mathbf{T} \in \mathbb{R}^{n \times m}$ correspond to the latent variables and constitute projections of the columns of $\mathbf{X}$ onto the basis $\{\mathbf{p_1}, \ldots, \mathbf{p_m}\}$.

To select the new basis $\mathbf{P}$, it is necessary to assess which characteristics of the transformed matrix $\mathbf{T}$ provide the optimal representation of the data. However, this evaluation can only be performed by making additional assumptions about the nature of the data:

PCA, Assumption 2. *There is a degree of linear dependence between variables, which hinders the description of the essence of the data due to the presence of redundant and unnecessary information*.

PCA, Assumption 3. *The data exhibit a high signal-to-noise ratio, indicating that larger variances carry greater relative importance of information*.

Hypothesis PCA-A2 indicates that it is desirable for the columns of $\mathbf{T}$ to be linearly independent. In other words, the empirical covariance matrix $\mathbf{C_T}$:

$$\mathbf{C_T} = \frac{(\mathbf{T} - \overline{\mathbf{T}})'(\mathbf{T} - \overline{\mathbf{T}})}{n-1}, \tag{2}$$

should only contain diagonal elements. In Equation (2), $\overline{\mathbf{T}}$ denotes the vector of column-wise means of $\mathbf{T}$. Before proceeding with the mathematical development, it is necessary to introduce one additional hypothesis:

PCA, Assumption 4. *Each variable in the data matrix has a zero mean, i.e.*, $\overline{\mathbf{X}} = \overline{\mathbf{T}} = 0$.

In practice, the validity of this assumption is ensured through a process called centralization, where the mean of each variable is subtracted. In this case, the covariance matrix of $\mathbf{T}$ can be expressed as:

$$\mathbf{C_T} = \frac{\mathbf{T}'\mathbf{T}}{n-1}. \tag{3}$$

Using Equation (1), Equation (3) can be rewritten in terms of the principal component matrix $\mathbf{P}$ as follows:

$$\mathbf{C_T} = \mathbf{P}'\mathbf{C_X}\mathbf{P}. \tag{4}$$

According to the Spectral Theorem [35], a symmetric matrix such as $\mathbf{C_X}$ can be decomposed into eigenvalues and eigenvectors as follows:

$$\mathbf{C_X} = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}', \tag{5}$$

where $\mathbf{E}$ is the matrix containing the eigenvectors of $\mathbf{C_X}$ as columns, and $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{C_X}$ in descending order of magnitude.

To obtain a diagonal matrix $\mathbf{C_T}$, the matrix $\mathbf{P}$ can be conveniently chosen as $\mathbf{P} = \mathbf{E}$. This is because substituting Equation (5) into Equation (4) yields:

$$\mathbf{C_T} = \boldsymbol{\Lambda}. \tag{6}$$

If Assumptions PCA-A2 and A3 are satisfied, selecting $\mathbf{P} = \mathbf{E}$ results in a transformed matrix $\mathbf{T}$ that is advantageous for data analysis in two ways. Firstly, it eliminates the redundancy caused by the linear dependence between the variables. Secondly, it organizes the variables in decreasing order of importance according to the variance values present in the diagonal matrix $\mathbf{C_T}$.

It is useful to interpret the PCA technique from an optimization perspective. In this context, the first principal component $\mathbf{p}_1$ should maximize the variance of the latent variable $\mathbf{t}_1 = \mathbf{X}\mathbf{p}_1$:

$$\begin{aligned}
\mathbf{p}_1 &= \underset{\|\mathbf{p}\|_2=1}{\arg\max} \left\{ (\mathbf{X}\mathbf{p})' \mathbf{X}\mathbf{p} \right\} \\
&= \underset{\|\mathbf{p}\|_2=1}{\arg\max} \left\{ \mathbf{p}'\mathbf{X}'\mathbf{X}\mathbf{p} \right\} \\
&= \underset{\|\mathbf{p}\|_2=1}{\arg\max} \left\{ \mathbf{p}'\mathbf{C_X}\mathbf{p} \right\} \\
&= \arg\max \left\{ \frac{\mathbf{p}'\mathbf{C_X}\mathbf{p}}{\mathbf{p}'\mathbf{p}} \right\}.
\end{aligned} \tag{7}$$

In the last equality, the unit norm constraint is relaxed by dividing the objective function by $\|\mathbf{p}\|_2^2$. The remaining principal components can be found in the successive orthogonal spaces by deflating the matrix $\mathbf{C_x}$ with respect to $\mathbf{p}$ [36].

The objective function represented by Equation (7) takes a mathematical form known as the Rayleigh quotient:

$$r(\mathbf{A}, \mathbf{B}, \mathbf{w}) = \frac{\mathbf{w}'\mathbf{A}\mathbf{w}}{\mathbf{w}'\mathbf{B}\mathbf{w}}. \tag{8}$$

It can be shown that the critical points of $r(\mathbf{A}, \mathbf{B}, \mathbf{w})$ correspond to the solution of the generalized eigenvalue problem [22,37]:

$$\mathbf{A}\mathbf{e} = \lambda\mathbf{B}\mathbf{e}. \tag{9}$$

Since $\mathbf{B} = \mathbf{I}$ in the objective function, Equation (9) simplifies to an eigenvalue problem, which is consistent with the derivation of Equation (6). It can be demonstrated that the eigenvalue decomposition of the covariance matrix $\mathbf{C_x}$ is equivalent to the singular value decomposition of the data matrix $\mathbf{X}$ [38]. In fact, the latter approach is computationally more efficient and stable, as it eliminates the need to calculate $\mathbf{C_x}$.

3.1.2. Dimensionality Reduction

In order to perform dimensionality reduction, a certain number a of principal components is selected such that they account for a certain percentage of the explained variance (usually 90%), resulting in a reduced matrix $\mathbf{P_a} \in \mathbb{R}^{m \times a}$. The reduced matrix of latent variables $\mathbf{T_a} \in \mathbb{R}^{n \times a}$ is obtained through the equation:

$$\mathbf{T_a} = \mathbf{X}\mathbf{P_a}. \tag{10}$$

Based on Assumption PCA-A3, the most informative variables are those that exhibit more variability in the multivariate space under study. Therefore, the matrix $\mathbf{T_a}$ would be a more effective and suitable representation of the data, as it excludes the variables associated with smaller variances (which, according to the assumption, correspond to measurement noise). This matrix can be projected onto the original variables as follows:

$$\hat{\mathbf{X}} = \mathbf{T_a}\mathbf{P_a}' = \mathbf{X}(\mathbf{P_a}\mathbf{P_a}'). \tag{11}$$

By doing so, the original variables can be reconstructed in the matrix $\hat{\mathbf{X}}$, which includes only the primary variation in the dataset, corresponding to the a leading principal components. In this sense, the PCA technique can be viewed as a filter, as the transformation from $\mathbf{X}$ to $\hat{\mathbf{X}}$, in theory, reduces the noise level of the dataset.

The difference between $\mathbf{X}$ and $\hat{\mathbf{X}}$ is represented by the residual matrix $\mathbf{E}$:

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}} = \mathbf{X}(\mathbf{I} - \mathbf{P_a}\mathbf{P_a}'). \tag{12}$$

The subspaces generated by $\hat{\mathbf{X}}$ and $\mathbf{E}$ are known as *principal space* and *residual space*, respectively, [20].

Several discussions regarding selection methods for the parameter a can be found in the literature [39–44]. In particular, Saccenti and Camacho [43] conducted a comparison of different techniques, including statistical-based methods, cross-validation methods and methods based on numerical approximation. The authors offered recommendations for selecting the most suitable methods based on the specific characteristics of the analyzed data.

Figure 1 illustrates the application of the PCA procedure to a bidimensional dataset. In the first plot, the red arrows depict the two principal components, with the larger arrow indicating the first principal component, associated with the direction of higher explained variance. The second plot demonstrates the linear transformation from $\mathbf{X}$ to $\mathbf{T}$. The third plot demonstrates the dimensionality reduction step, where the bidimensional data is transformed into a unidimensional representation. Finally, the last plot presents the reconstructed data in the original bidimensional space, now with variation only along the direction of the principal component.

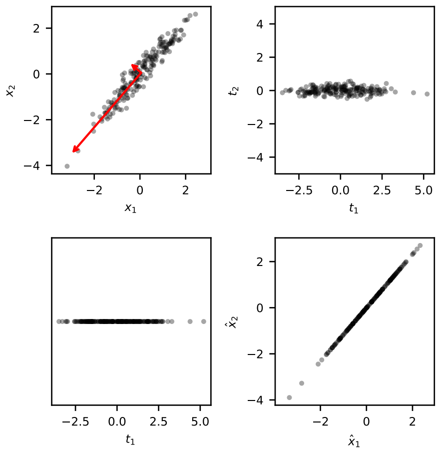

Figure 1. Schematic illustration of the PCA procedure: the identification of principal components (depicted by red arrows), the subsequent projection to a latent space, the dimensionality reduction step and the resulting reconstruction in the original space.

3.1.3. Process Monitoring

In the context of process monitoring, the data matrix $\mathbf{X}$ is often derived from a time series. The aim of applying the PCA technique in this scenario is to identify the structure of the principal subspace during a normal operating period and use it as a reference to evaluate new observations in order to determine whether the structure is maintained.

When the process that generates the time series exhibits dynamic behavior, i.e., temporal dependence, the PCA technique may not be able to accurately identify the structure of the principal subspace. In such cases, the latent variables may be autocorrelated and potentially exhibit cross-correlation [45]. Hence, to utilize PCA for process monitoring purposes, the following assumption needs to be considered:

PCA, Assumption 5. *The data matrix is obtained from a steady-state process.*

The monitoring procedure utilizing PCA can be carried out in two distinct stages. The first stage, referred to as training, involves applying PCA to a dataset collected from normal operation, which is commonly known as the training set. The outcome of the training process includes two matrices: the projection matrix $\mathbf{P_a}$ and the explained variances matrix $\mathbf{\Lambda_a}$. In the second stage, known as testing, new observations $\mathbf{x}$ are subjected to monitoring using statistical indices. The most frequently employed indices for monitoring are the T^2 and Q indices, which are described below.

In the context of PCA, the T^2 statistic is computed by applying the confidence ellipse of the normal distribution to the latent variable $\mathbf{t_a}$, which is the projection of the observation vector $\mathbf{x}$ onto the first a principal components [27,46]:

$$T^2 = \mathbf{t}'_\mathbf{a}\mathbf{\Lambda_a}^{-1}\mathbf{t_a} = \mathbf{x}'\mathbf{P_a}\mathbf{\Lambda_a}^{-1}\mathbf{P}'_\mathbf{a}.\mathbf{x} \quad (13)$$

The T^2 statistic provides a global measure of the variability captured by the PCA model in the first a principal components, which is associated with the systematic variations in the process being monitored [20]. During the monitoring phase, an increase in the T^2 value for new observations indicates the occurrence of disturbances with a similar nature to the variations identified by the PCA model, that is, disturbances in the directions of the plane described by the model [47].

The Q statistic is defined as the Euclidean distance between an observation vector $\mathbf{x}$ and its projection onto the principal subspace $\hat{\mathbf{x}} = (\mathbf{P_a}\mathbf{P}'_\mathbf{a})\mathbf{x}$:

$$Q = ||\mathbf{x} - \hat{\mathbf{x}}||^2 = ||(\mathbf{I} - \mathbf{P_a}\mathbf{P}'_\mathbf{a})\mathbf{x}||^2 = (\mathbf{x} - \hat{\mathbf{x}})'(\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{x}'(\mathbf{I} - \mathbf{P_a}\mathbf{P}'_\mathbf{a})\mathbf{x}. \quad (14)$$

The Q statistic is the squared norm of the prediction errors (residuals) of the PCA model and is often referred to as SPE (square prediction error). It provides a global measure of variability that is not captured by the PCA model, i.e., variability present in the residual space, which is associated with unusual or unexplained variations in the process being monitored [20]. During monitoring, an increase in the Q value for new observations suggests a change in the relationships between the variables, such as a violation in correlations, resulting in a break in the structure captured by the model during training. In other words, the plane described by the PCA model does not include the direction of the disturbance [47]. Due to the form of Equation (14), establishing a detection limit using the Q statistic is equivalent to delimiting a hypersphere as a confidence region in the residual space [48].

The PCA procedure is illustrated in the present paper through the application of the Tennessee Eastman Process (TEP) dataset. This dataset is widely used in the process monitoring literature and originates from a chemical process simulation developed by Eastman Chemical Company [49]. The TEP comprises five units: reactor, condenser, vapor–liquid separator, stripper, and recycle centrifugal compressor. It involves four reactants and one inert being fed into a reactor, which generates two products and one byproduct through exothermic and irreversible reactions. For a comprehensive critical analysis of the TEP dataset, readers are referred to Melo et al. [5].

Figure 2 displays the application of the PCA training procedure to the TEP dataset. The first plot shows the growth in accumulated explained variance as the number of components increases. The second plot presents the control chart for the Q index when normal data are used.

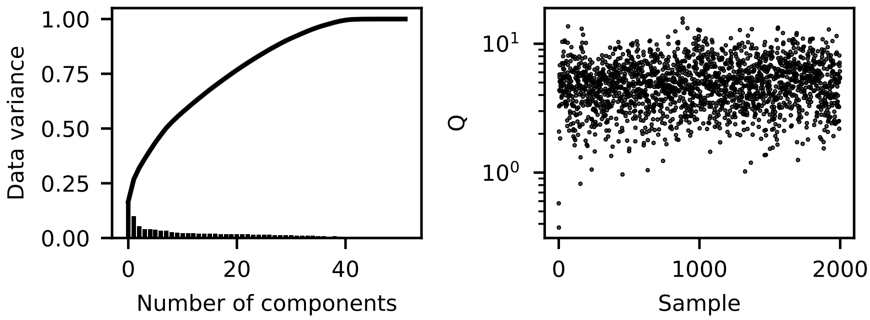

Figure 2. Illustration of the PCA training procedure for process monitoring: explained variance and Q control chart.

The use of the T^2 statistic, which is derived from the normal distribution, requires the formulation of an additional hypothesis:

PCA, Assumption 6. *The dataset* $\mathbf{X}$ *follows a normal distribution.*

In some cases, the T^2 statistic can be used even when Assumption PCA-A6 is violated, since the latent variables are linear combinations of the original variables. In this case, one may assume that the latent variables approximately follow a normal distribution due to the Central Limit Theorem [50]. This theorem establishes that the sum of random variables asymptotically approaches a normal distribution as the number of variables increases.

Some references consider normality as an intrinsic requirement for applying the PCA technique and not as a requirement for the use of the T^2 statistic, as considered in the present text. This happens because these references apply and interpret PCA in the context of statistical (in)dependence, while in the present text, the discussion involved the less rigorous concept of linear (in)dependence (see Assumption PCA-A2). To ensure that linear independence between the latent variables is equivalent to statistical independence, some hypotheses about the probability distribution of the data must be enunciated. It can be shown that assuming that the data follows a normal distribution is a sufficient condition to guarantee this equivalence [51].

This issue is subtle and often a source of conceptual confusion. For instance, Jonathon Shlens included the normality assumption in the early versions of his excellent tutorial on PCA [52] but removed it in later versions [38]. To support the understanding of this topic, it is worth consulting a concise and elucidative comparison between the terms linear independence, orthogonality, and uncorrelatedness presented by Rodgers et al. [53].

The Q and T^2 statistics are applicable to the two primary activities conventionally associated with process monitoring, namely fault detection and diagnosis.

3.1.4. Fault Detection

To carry out fault detection, control limits must be established for the statistics T^2 and Q. Considering a confidence level α, the most commonly used limits are [47,54]:

$$T^2_\alpha = \frac{a(n^2-1)}{n(n-a)} F_{\alpha(a,n-a)}, \tag{15}$$

$$Q_\alpha = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0}, \tag{16}$$

where:

- $F_{\alpha(a,n-a)}$ is the upper limit of the α percentile of the Fisher-Snedecor F distribution, with a and $n-a$ degrees of freedom;
- $\theta_i = \sum_{j=a+1}^{m} \lambda_j^i$;
- $h_0 = 1 - \left(\frac{2\theta_1\theta_3}{3\theta_2^2} \right)$;
- $c_\alpha = 1.645$ is the z-score corresponding to the $(1-\alpha)$ percentile of the standard normal distribution.

Equation (15) is derived directly from statistical theory, while Equation (16) is based on the approximation of the distribution of quadratic forms and was first proposed by Jackson and Mudholkar [55]. When the underlying normality assumptions are not met, Equations (15) and (16) may not perform well, and alternative control limits based on specific statistical properties of each dataset need to be developed.

Figure 3 presents the Q control chart for one of the faults in the Tennessee Eastman Process dataset. The fault event is indicated by the vertical line at sample 160. Following the occurrence of the fault, the data points exceed the horizontal dashed line, which represents the Q_α control limit. This triggers an alarm, indicating a deviation from the expected behavior of the process.

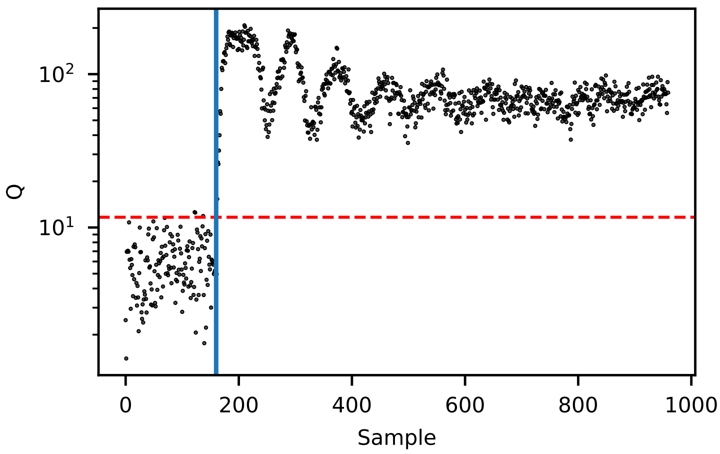

Figure 3. Illustration of the PCA Q control chart for a fault of the Tennessee Eastman Process dataset.

3.1.5. Fault Identification and Diagnosis

There exists a certain degree of confusion in the process monitoring literature regarding the precise definitions of the terms identification and diagnosis. For instance, Chiang et al. [20] defined identification as the act of indicating the most relevant variables for performing diagnosis, which was in turn defined as the determination of the root cause of the abnormal state. In contrast, Qin [18] proposed a different definition, in which the meanings of the terms are reversed: identification was defined as the process of identifying a fault from a set of possible known faults, while diagnosis referred to the identification of the variables that contributed to the occurrence of the fault. The latter definition is adopted in the present text.

There are several diagnostic methods in the context of PCA, among which the most popular are the contribution methods [20,50,56–58]. These methods are based on analyzing the contribution of each variable to the detection index. In order to apply these methods, an additional hypothesis must be assumed:

PCA, Assumption 7. *At the time of fault, the magnitude of the contribution of a variable to the increase in the detection rate is proportional to the influence of the variable on the occurrence of the abnormal event.*

The contribution methods used in PCA can be broadly classified into three categories [56]:

- Decomposition contributions: These methods decompose the detection index into the sum of contributions from individual variables.
- Reconstruction-based contributions: These methods consider the contribution of a variable to be the magnitude of the reconstruction of the detection index along the direction of that variable.
- Diagonal contributions: In multiblock problems, similar variables are grouped into blocks based on domain knowledge, which simplifies the scenarios with many variables. Diagonal contributions keep only the block terms in the detection index.

The formulation of the most widespread technique, the method of decomposition contributions, will be presented as an illustration. The indices T^2 and Q can be written in the form of a generalized index, idx:

$$idx = \mathbf{x}'\mathbf{M}\mathbf{x}, \tag{17}$$

where:

- $\mathbf{M} = \mathbf{P_a}\mathbf{\Lambda_a}^{-1}\mathbf{P}'_\mathbf{a}$ for $idx = T^2$;
- $\mathbf{M} = \mathbf{I} - \mathbf{P_a}\mathbf{P}'_\mathbf{a}$ for $idx = Q$.

The decomposition contribution of variable i to the generalized index is denoted as cd_i and expressed as [56]:

$$cd_i = \mathbf{x}'\mathbf{M}^{1-\beta}\mathbf{I_i}\mathbf{I}'_i\mathbf{M}^{\beta}\mathbf{x}, \tag{18}$$

where $\mathbf{I}_i$ constitutes the ith column of the identity matrix and $0 \leq \beta \leq 1$. According to the choice of the β parameter, different degrees of decomposition can be defined. In the literature, $\beta = 0$ or $\beta = 1$ are used under the name "partial contribution" and $\beta = 1/2$ is used under the name "complete contribution".

Figure 4 displays a partial contribution heatmap for a fault in the Tennessee Eastman Process dataset. This heatmap provides insights into the variables that contribute to the detection index at each moment in time, enabling the diagnosis of the variables that are associated with the evolution of the fault.

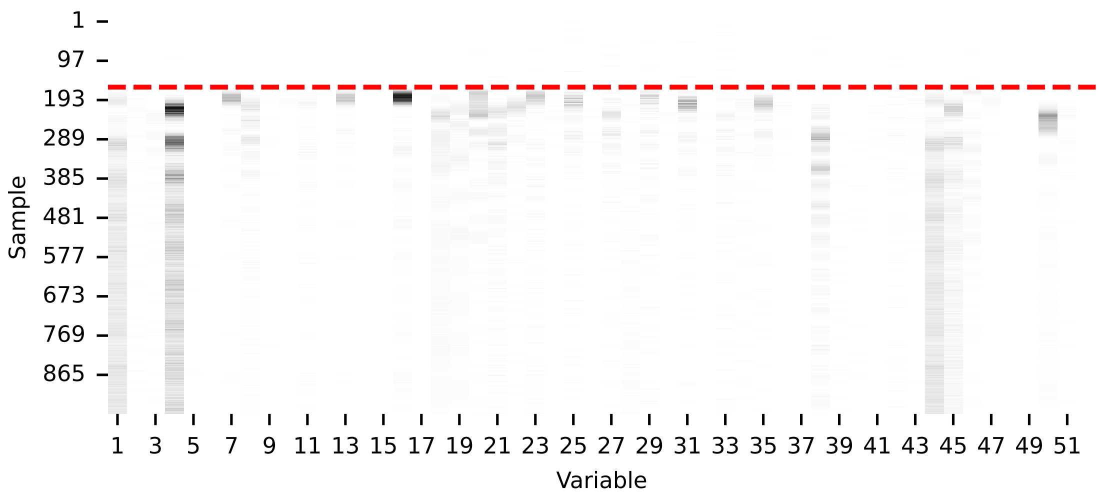

Figure 4. Illustration of the PCA diagnosis partial contributions for a fault of the Tennessee Eastman Process dataset. The red dashed line indicates the occurrence of the fault.

Despite their widespread use, the contribution methods mentioned above may suffer from a problem known in the literature as the smearing-out effect. This occurs when contributions from variables affected by the fault increase contributions from unaffected variables [59]. A detailed analysis of this effect was reported by Kerkhof et al. [60], who concluded that it is more pronounced along the principal components. The authors recommended applying contribution methods only in cases where there is a clear physical interpretation of the principal components. If such an interpretation is not available, they suggest using the deviation of each variable in relation to the expected value as an alternative for diagnosis (this procedure is equivalent to performing univariate diagnosis). In this context, a methodology to address the smearing-out effect was recently proposed by Alakent [61], which involves the application of a non-linear dimensionality reduction method.

Figure 5 illustrates the flowchart of the complete process monitoring pipeline described in this section, including the following steps: data collection, model training, threshold setting, fault detection, fault diagnosis, fault identification, and process recovery.

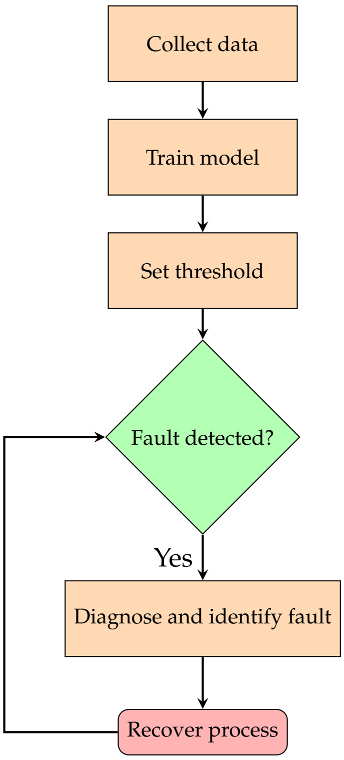

Figure 5. Process monitoring flowchart.

3.1.6. Evolution and Recent Trends

Among the multivariate monitoring techniques, PCA has received the most attention in both simulated and real systems. Nevertheless, in its fundamental form, as outlined in earlier sections, it may not be adequate for a multitude of significant industrial applications. As a result, there have been several suggestions for enhancing its efficacy, which can be summarized as follows.

- **Batch processes**: Since the early 1990s, when PCA was first applied, adaptations have been made to utilize the model for batch processes through the implementation of multi-way PCA [62,63]. This technique transforms the data matrix of dimension $n \times m \times b$ (where the third dimension b refers to multiple batches) into a two-dimensional matrix through dimension reduction, which can be achieved either by time or by batch [18]. To compare the effectiveness of monitoring batch processes using multi-way PCA with other strategies such as parallel factor analysis and trilinear decomposition, it is recommended to refer to comparison articles [64–66], as well as review articles [67,68]. A promising technique that has not been discussed in these reviews is 2D-PCA [69,70], which enables the separate monitoring of the dynamics associated with variables and batches.

- **Dynamics**: When the process being analyzed does not meet the assumption of being in a steady state, exhibiting dynamic characteristics such as temporal dependence and autocorrelation, the confidence regions from the PCA monitoring statistics become too large, resulting in high rates of undetected faults [18]. To address this problem, the data matrix can be expanded by incorporating the lagged values of the original variables [45]. However, this approach increases the number of parameters and principal components, which contradicts the dimensionality reduction benefit of PCA, lacks an explicit representation of the relationship between latent and original variables, and fails to establish a connection between latent variables and the dynamic content of the data [71]. To mitigate these drawbacks, recent methods have been proposed to explicitly create dynamic principal components [71–73]. For example, DiPCA (dynamic-inner PCA) [71,74] separates the latent variables into two groups: a set of principal time series and a residual with little or no autocovariance. The latter group can be treated as a static dataset by traditional methods, such as PCA. DiPCA was recently applied to assemble a sequential structure for extracting and modeling dynamic and non-linear characteristics of data [75] and to solve the time series segmentation problem [76], which involves generating subsequences based on changes in the dynamics. Adaptive algorithms are another strategy to deal with dynamics by automatically adjusting to changes in mean, variance, and/or correlation structures in the data [18]. This strategy is especially useful for processes with drifts or changes in operating points. Examples of such algorithms include RPCA (recursive PCA), MW-PCA (moving window PCA), and EWMA-PCA (exponentially weighted moving average PCA) [46,77–80]. A review of recent advances in monitoring dynamic processes was presented by Aldrich [81].
- **Nonlinearity**: There are two main strategies for incorporating nonlinearity into PCA. The first involves the use of neural networks, as proposed by several researchers: (i) Kramer [82], who developed an approach based on associative networks; (ii) Dong and McAvoy [83], who combined neural networks using the principal curves algorithm [84]; (iii) Jia et al. [85], who applied a network called IT-net [86] to process monitoring; and (iv) Liu et al. [87], who combined networks with the SVDD (support vector data description) technique, a machine learning method described in Section 4. The second strategy is the use of the kernel-PCA (KPCA) technique, developed by Scholkopf et al. [88], applied in process monitoring by Lee et al. [89] and Choi et al. [90] and further developed in several works [91–98]. A kernel is a function that maps data points into a higher-dimensional space, where nonlinear representations can be captured using linear relationships [99]. In addition to these two main strategies, there are other approaches for adapting PCA to nonlinearity, such as the use of the genetic algorithm [100]. Recent reviews on applications of the kernel technique in process monitoring have been published [99,101,102].
- **Multimodality**: To address multimodality, clustering algorithms can be utilized to distinguish between operation modes [103]. One straightforward approach involves identifying clusters based on prior knowledge about the process or specialized algorithms. Subsequently, individual PCA models are constructed for each cluster [18]. Another effective approach is the application of Gaussian mixture models, as proposed in several works [104–108], which enable the automatic identification of clusters during the PCA training process. A comprehensive review of multimodal process monitoring has been presented by Quiñones-Grueiro et al. [103].
- **Multiple time scales**: Bakshi et al. [109] conducted the pioneering work of analyzing process data across different time scales, utilizing the methodology proposed by Rioul and Vetterli [110], which involves the application of wavelets for data processing. Kosanovich and Piovoso [111,112] used this approach to pre-process data before

applying PCA for monitoring. Bakshi [113] integrated PCA and wavelets into a single method, called MSPCA (multiscale PCA), which works by generating a PCA model for each of the time scales of interest and then joining the models to create a single multiscale model. Since this work, several applications have been reported using or adapting this approach, involving: extension to the diagnostic problem [114,115], theoretical analysis [116], generation of contribution plots [117], analysis of data with multiresolution characteristics [118], and combination with the kernel technique for modeling nonlinearity [119]. Recent works include the application in multi-unit and batch processes [120], the integration with the generalized likelihood ratio test [121] and the use in conjunction with kernel techniques to deal with non-linearities [122]. Extensive reviews on the application of multiscale methods in process data analysis have been published [123,124].

- **Multiple units and/or stages**: To address the issue of multiple units and/or stages in a decentralized manner, multiblock methods such as CPCA (consensus PCA) and HPCA (hierarchical PCA) have gained widespread use [18]. These techniques involve dividing the data matrix into several blocks simultaneously to obtain local and global information. An important contribution was the demonstration of the algorithmic equivalence between regular and multiblock methods [125]. Recently, fully data-driven approaches for partitioning units/stages have gained attention [126]. Notable examples of these approaches include: the proposal of constructing sub-blocks along the principal components [127]; the combination of spectral clustering based on mutual information, Bayesian inference, and KPCA to handle non-linearity [128]; the combination of KPCA with SVDD to determine the independent variables of the process [126]; and the development of a hybrid method for variable categorization [129]. Comprehensive reviews on plant-wide/distributed monitoring have been published [130,131].
- **Error handling**: When errors are not independent and identically distributed, incorporating information on the probability density of variables can enhance the modeling of the residual space and improve effectiveness [132,133]. Probabilistic PCA [134], Bayesian PCA [135], and MLPCA (maximum likelihood PCA) [90,136–139] are some of the approaches used to address this issue. It is noteworthy that MLPCA [132] or principal component regression [140] can handle heteroskedastic data (i.e., data with variance dependent on time or operating conditions). An interesting proposal in this regard is the integration of PCA methodology with data reconciliation [141]. Ge [142] provided a tutorial review on the application of probabilistic modeling on latent variables for process data analysis.
- **Poor data quality**: To address problems related to poor data quality, such as missing or spurious observations, a class of techniques called robust PCA has been developed. Pioneering work by Walczak and Massart [143] and Xie et al. [144] laid the foundation for this field, which was further developed with applications and extensions [145–151]. It is interesting to note that robust PCA is widely used in computer vision, image processing, and video monitoring applications [152]. Severson et al. [153] published a review of PCA applications to process data with missing observations. Zhu et al. [154] provided an informative discussion comparing PCA to other robust methods from a big data perspective.

3.2. PLS/CCA

3.3. Fundamentals

PLS (partial least squares) and CCA (canonical correlation analysis) are two techniques with similar objectives, which are to generate two sets of linear combinations of the original variables (one for each category of variables, predictor and predicted) with the goal of selecting the combinations that best summarize the linear relationship between the two categories [22,36,37]. CCA was introduced by Hotelling [155], while PLS was first proposed by Wold [156]. Although PLS is widely used in chemometrics, which may explain its prevalence in process monitoring research, the preference for PLS over CCA in chemometrics has been recently challenged [157]. It is worth noting that PLS is also popular in other fields, such as the biopharmaceutical sector, where it is the most commonly used data-driven model [158].

The following assumptions about the nature of the data are shared by both techniques:

PLS/CCA, Assumption 1. *The latent variables can be obtained by performing a linear transformation on the dataset.*

PLS/CCA, Assumption 2. *The dataset can be pre-classified into two distinct categories, namely the predictor set* $\mathbf{X}$ *and the predicted set* $\mathbf{Y}$*.*

PLS/CCA, Assumption 3. *Within each set* $\mathbf{X}$ *and* $\mathbf{Y}$*, there is a degree of linear dependence among the variables, which hinders the description of the essence of the data due to the presence of redundant and unnecessary information.*

PLS/CCA, Assumption 4. *There is a latent linear relationship between the set* $\mathbf{X}$ *and the set* $\mathbf{Y}$*, which is of interest to be made explicit.*

The difference between PLS and CCA lies in the way of obtaining the latent linear relationship mentioned in PLS/CCA-A4: while PLS seeks the directions that maximize the covariance between $\mathbf{X}$ and $\mathbf{Y}$, CCA seeks the directions that maximize the correlation between $\mathbf{X}$ and $\mathbf{Y}$ [22].

PLS, Assumption 5. *The latent linear relationship stated in Assumption 4 is better represented by the directions that maximize the* covariance *between* $\mathbf{X}$ *and* $\mathbf{Y}$*.*

CCA, Assumption 5. *The latent linear relationship stated in Assumption 4 is better represented by the directions that maximize the* correlation *between* $\mathbf{X}$ *and* $\mathbf{Y}$*.*

Finally, another shared assumption is that it is necessary for the mathematical development of both techniques:

PLS/CCA, Assumption 6. *Each variable in the data matrices has a zero mean, i.e.,* $\overline{\mathbf{X}} = \overline{\mathbf{Y}} = 0$*.*

Just like PCA, PLS and CCA problems can be formulated as the maximization of the Rayleigh quotient.

For PLS:

$$\mathbf{p}_{1,x}, \mathbf{p}_{1,y} = \underset{\|\mathbf{p}\|_2=1}{\arg\max} \left\{ \frac{\mathbf{p}_x^T \mathbf{C}_{xy} \mathbf{p}_y}{\sqrt{\mathbf{p}_x^T \mathbf{p}_x} \sqrt{\mathbf{p}_y^T \mathbf{p}_y}} \right\}. \tag{19}$$

For CCA:

$$\mathbf{p}_{1,x}, \mathbf{p}_{1,y} = \underset{\|\mathbf{p}\|_2=1}{\arg\max} \left\{ \frac{\mathbf{p}_x^T \mathbf{C}_{xy} \mathbf{p}_y}{\sqrt{\mathbf{p}_x^T \mathbf{C}_x \mathbf{p}_x} \sqrt{\mathbf{p}_y^T \mathbf{C}_y \mathbf{p}_y}} \right\}. \tag{20}$$

This allows the problems to be solved using the framework of generalized eigenvalue problems, as presented in Equation (9) [37]. The monitoring statistics T^2 and Q are formulated similarly to those of PCA [20,131].

There is currently no consensus on the appropriate situations to use partial least squares (PLS) or canonical correlation analysis (CCA). Some studies suggest using CCA to model input–output relationships and PLS to monitor quality variables, which are typically more challenging to measure than process variables, resulting in issues such as delays [131,159,160]. Zhang et al. [160] also noted that CCA is typically applied at a local, process-specific level, whereas PCA and PLS are applied at a global level, encompassing the entire plant. In the field of computer science, Taylor and Cristianini [36] argued that PLS is appropriate for establishing relationships between input and output representations, while CCA is suitable for finding relationships between different representations of the same dataset. Another approach for comparing PLS and CCA involves investigating their mathematical relationship. Barker and Rayens [161], for instance, view PLS as a penalized CCA, where the penalties are represented by a PCA model in the space $\mathbf{X}$ and a PCA model in the space $\mathbf{Y}$.

Confusion in the literature exists regarding the nomenclature of these techniques. For example, the acronym PLS can have two meanings: the original name, partial least squares, or a reinterpretation, projection to latent structures [162]. Moreover, there are many variants of PLS [163], and it is often unclear which variant is being referred to when using the acronym. The term canonical correlation analysis (CCA) is used synonymously with canonical variate analysis (CVA) in some references [20,164]. However, CVA is better understood and accepted in the literature as the application of CCA methodology for identifying subspaces to find a linear combination of past inputs and outputs that is more predictive of future outputs [165,166]. Therefore, CVA has a more restricted meaning since nothing in the CCA technique mathematically imposes its application in this specific context. The confusion increases when analyzing literature from other domains. For example, in a book on ecology, Gittins [167] defines CVA as an application of CCA in which variables with arbitrary values are added to the dataset, to which the analyst assigns meanings artificially to include some data characteristic in the analysis (such as information on classes). This definition of the term is entirely unrelated to what is discussed in the fields of chemometrics and process monitoring.

The history and trends of PLS/CCA techniques, especially PLS, follow the same lines as the development of PCA. Works have been developed for applications concerning batch processes [168–174], dynamics [169,175–192], nonlinearity [119,185,193–198], multimodality [199–202], multiple time scales [119,203–205], multiple units and/or stages [125,170,193,206–213], error handling [214–217] and poor data quality [216,218–222]. In addition, new topics have emerged, such as the study of the effect of dataset size on model structure [223]. For a comprehensive review of the CCA technique, the reader is referred to Yang et al. [224].

3.4. Other Techniques

As previously mentioned, PCA and PLS/CCA are multivariate statistical techniques widely used in the field of process monitoring. However, it is important to note that there are other alternative methodologies available, which are worth considering. Some of these methodologies include:

- Fisher discriminant analysis (FDA): Introduced by Fisher [225], this technique assumes the a priori division of variables into two categories and aims to generate an orthogonal set of linear combinations of the original variables that effectively separate these categories. FDA is particularly valuable when data is available for both normal operation and multiple fault cases [18]. Various examples of FDA applications in monitoring problems have been proposed [226–240]. For more detailed information on this technique, please refer to Chiang et al. [20].
- Independent component analysis (ICA): Introduced by Jutten and Herault [241], ICA generates a set of linear combinations of the original variables to maximize the independence statistic between these combinations. ICA is considered an alternative to PCA when the data distribution is not assumed to be normal [242]. The application of ICA in process monitoring is relatively recent, with the initial works proposed in the early 2000s [243–248]. Subsequent advancements have been proposed in the recent literature [249–259]. Notably, Zhang et al. [260] proposed an interesting method that automatically selects between the exclusive or mixed application of PCA and ICA techniques, with or without using kernel functions to incorporate nonlinearity. For more detailed information on the technique, please refer to Tharwat [261] or Comon [262]. For a comprehensive review of ICA's application in process monitoring, it is recommended to consult Palla and Pani [263].

- Slow feature analysis (SFA): Proposed by Wiskott and Sejnowski [264], this technique is based on a non-linear optimization formulation that aims to extract a set of independent variables that exhibit slow variations over time. The underlying hypothesis is that slow variations are more relevant and informative compared to fast variations. SFA has gained significant popularity in the recent literature on process monitoring [265–282]. For more detailed information on the technique, please refer to Song and Zhao [283].

4. Machine Learning

Artificial intelligence (AI) is a rapidly evolving interdisciplinary field of science and engineering that has gained significant prominence in recent years. It is considered one of the most promising domains, alongside biotechnology, in terms of potential advancements, innovation, and societal impact [284–286]. The term "artificial intelligence" was first introduced in 1956 by J. McCarthy, M. L. Minsky, H. Simon, and A. Newell as:

> "*(...) the ability of machines to understand, think, and learn in a similar way to human beings*" [286].

Kaplan and Haenlein [287] proposed a more specific definition:

> "*We define AI as a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation*" [287].

From these definitions, the significance of the learning process within AI becomes apparent. The subfield of AI that focuses on learning is commonly referred to as machine learning. Although the terms are occasionally used interchangeably, it is essential to emphasize that AI extends beyond machine learning. It encompasses various aspects, including data perception, knowledge representation, and the implementation of intelligent devices [284,287]. For an extensive discussion on the impact of AI in chemical engineering, we recommend reading Venkatasubramanian [285]. Regarding the specific impact of machine learning, several works have been published [16,288–295].

To gain a deeper understanding of the nature of machine learning methods, it is beneficial to examine some proposed definitions of the term [296]:

> "*A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E*" [297].

> "*Machine learning is essentially a form of applied statistics with increased emphasis on the use of computers to statistically estimate complicated functions and a decreased emphasis on proving confidence intervals around these functions*" [298].

The first definition, which takes a pragmatic approach, focuses on understanding the concept of "learning" within the term "machine learning". It highlights the operational aspects of running a computer program in a manner that can be considered learning. The second definition, taking a more theoretical perspective, provides a framework to visualize the relationship between machine learning and statistics. In this interpretation, machine learning is seen as an application of statistics, with the primary difference lying in the approach used. The use of the term "increased" is particularly intriguing as it captures a historical trend in the definition. As computational processing power continues to advance, learning algorithms increasingly rely on this capacity to estimate unknown functions, often sacrificing the interpretability of the underlying mechanisms and results for humans. This trend has witnessed significant growth, particularly with the emergence of deep neural network models [285,298].

Machine learning techniques are primarily classified based on the type of learning they employ. Supervised learning involves training models using a labeled training dataset, where each sample is associated with a desired solution or outcome. Conversely, unsupervised learning involves training models on an unlabeled training dataset, where the desired solution is unknown. Semi-supervised learning combines aspects of both supervised and unsupervised learning, as it utilizes a training dataset containing both labeled and unlabeled data. Another significant type of learning is reinforcement learning, in which an algorithm interacts with an environment by observing its state, selecting and executing actions, and subsequently receiving rewards or penalties. By exploring the environment and adjusting actions according to its consequences, the algorithm aims to maximize the cumulative reward it obtains over time. Supervised and unsupervised learning algorithms are widely used in process monitoring applications, while semi-supervised learning is gaining traction [230,299–301]. Within the context of PSE, reinforcement learning is predominantly used in automatic process control applications [302–305], although a few applications in process monitoring have been proposed [306–308].

The application of approaches from the machine learning literature to process monitoring problems can be traced back to the late 1980s when neural networks were first utilized [309–311]. However, during that time, such applications were limited to a few studies, with the predominant focus in the field centered around latent variable models from the perspective of multivariate statistical theory. In recent years, there has been a notable shift in the process monitoring community, driven by the development of data science and the successful application of machine learning methods, particularly deep learning, across various domains. This shift has resulted in an increasing emphasis on data science and machine learning approaches, taking precedence over traditional multivariate statistics, starting in the 2010s.

The analysis of key works by Joe Qin, a prominent researcher in the field, provides a clear illustration of this historical evolution. His important reviews on process monitoring techniques [14,18,25] primarily focus on latent variable models from the perspective of multivariate statistics. In the 2003 article [14], there is no mention of the term "machine learning". In the 2009 article [25], there is only one vague reference to the term, appearing in the second line of the article, with a contextualizing meaning. By the 2012 article [18], the term "machine learning" starts to gain prominence, although it is presented as an alternative approach to the main methodology being discussed:

> "*A related category of nonlinear data-driven methods suitable for process monitoring is from the machine learning literature including neural networks [310] and support vector machines (SVM) [312]. Work in this area is rich by itself and a significant amount of work related to process monitoring is reviewed in Venkatasubramanian et al. [313]*" [18].

It is noteworthy that in an article titled "Survey on data-driven industrial process monitoring and diagnosis," the author allocates thirteen pages to methodologies related to PCA and PLS, while dedicating only one paragraph to machine learning techniques. Furthermore, the author directed readers seeking more detailed information on machine learning to a review paper that was nearly a decade old at the time of publication.

Two years later, in a perspective article [19], more attention was directed towards machine learning methods in the context of PSE. The article highlights the differences between machine learning and multivariate statistics, acknowledging the need to incorporate advances achieved in other areas into PSE:

> "*Although multivariate statistical approaches have been the favorite choice in the most recent 2 decades, it should be noted that data have been an integral part of process engineering solutions since time models have been used for process optimization and control. (...) However, these aforementioned data analytics and practice in process systems engineering have apparently not connected to the recent development in machine learning, data mining, and big-data analytics. They differ not only in terms of sizes but also in how and what data should be used in solving real operation problems. In some ways, process systems engineering solutions are confined to one set of principles that are believed to be sound, whereas the machine learning and data-mining communities take the other way and achieve unexpected results and solutions that defy conventional wisdom*" [19].

In the most recent work [16], a noticeable shift in discourse is observed. Machine learning has taken center stage in the discussion, as evident from the title, "Advances and opportunities in machine learning for process data analytics". Latent variable models, which were once the core focus, are now categorized as part of the suite of machine learning techniques:

> "*Although we consider multivariate latent variable methods as a subset of the statistical learning methods, the field of machine learning has grown tremendously in the last two decades* (...) *Depending on the characteristics of data and models, such as linear, collinear, or nonlinear data, there are respective supervised and unsupervised learning methods to choose from. To deal with collinearity in the data, for example, PCA is unsupervised learning and PLS is supervised learning*" [16].

Considering multivariate statistical models as a subset of machine learning techniques is a conceptually elegant approach and holds theoretical validity. However, in the current study, we have made a deliberate choice to maintain rigor in describing the historical development. Consequently, it was considered more appropriate to present the techniques in separate sections.

In the subsequent portion of this section, an extensive overview of neural network models, considered one of the most successful models in machine learning, is presented from a historical and didactic standpoint, specifically in the context of process monitoring. Following that, a concise summary and review of other methods will be provided.

4.1. Artificial Neural Networks

4.1.1. Fundamentals

Artificial neural networks constitute a category of learning techniques that draw loose inspiration from the functioning of the human brain [314,315]. In the brain, neurons are interconnected through synapses, which act as connections between dendrites (inputs) and axons (outputs). These synapses undergo modulations in response to external stimuli, and it is through these modulations that the learning process occurs in living organisms [315].

Figure 6 illustrates the architecture of an artificial neuron. Within the network, each neuron k acts as a local processing unit that receives n inputs ($x_1, x_2, \ldots, x_n$) and generates an output (y_k). The connections between the inputs and the neuron are assigned weights, represented as w_{ik} for each connection i. The term b_k is referred to as the bias and serves as a parameter that is independent of the inputs. It is analogous to the intercept in a linear regression model.

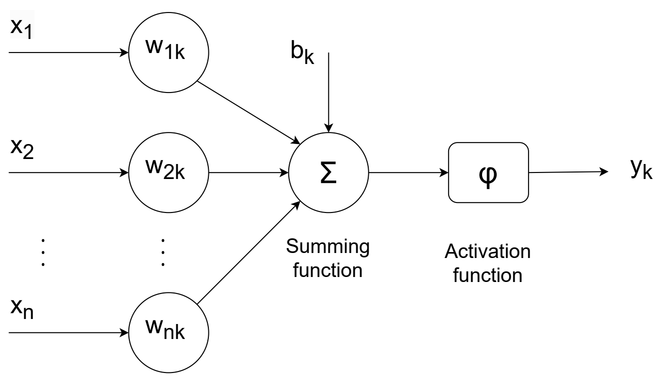

Figure 6. Artificial neuron model.

The output (y_k) of the neuron can be expressed as follows:

$$y_k = \phi\left(\sum_{i=1}^{n} w_{ik}x_i + b_k\right), \tag{21}$$

where ϕ is a non-linear activation function applied to the weighted sum of its inputs plus the bias term. In the neural network, the connections between multiple neurons form the network architecture. Similar to the learning process in living organisms, learning in this case occurs by adjusting the weights associated with the connections. There are several algorithms used for training neural networks, which can be categorized according to four basic learning rules, enumerated below [316]:

- **Error correction learning (delta rule)**: in this type of supervised algorithm, the weights are adjusted to minimize the network prediction error, which consists of the difference between the network output and the expected result.
- **Competitive learning**: In this unsupervised algorithm, the weights are adjusted through competitive interactions between neurons. The objective of each neuron is to achieve the highest representation power relative to the patterns in the learning data.
- **Boltzmann learning**: In this unsupervised algorithm, the weights are adjusted based on the principle of energy minimization. The goal is to capture statistical dependencies in the data and create a model capable of generating samples similar to the training set.
- **Hebbian learning**: In this unsupervised learning method, weights are adjusted based on Hebb's principle. According to this principle, if neurons downstream and upstream of a connection are activated simultaneously, the connection should be strengthened. Conversely, if the activation is asynchronous, the connection should be weakened.

Among the four types of learning mentioned, error correction learning is the most extensively applied and will be the primary focus of this section. In this context, the perceptrons are the simplest neurons capable of making decisions. A perceptron is a type of neuron that employs the binary step function as its activation function, defined as:

$$f(u) = \begin{cases} 0, & \text{if } u \leq 0 \\ 1, & \text{if } u > 0. \end{cases} \quad (22)$$

By utilizing this activation function, it is possible to train the weights (w) of the neuron to create a binary classifier. To accomplish this, a straightforward algorithm known as the perceptron learning rule is employed. The perceptron learning rule guarantees convergence for cases where linear separability exists, and further details can be found in the work by Russell and Norvig [284].

To introduce greater complexity and flexibility in classification tasks, perceptrons can be connected in a network configuration, as depicted in Figure 7, which illustrates a network with two hidden layers.

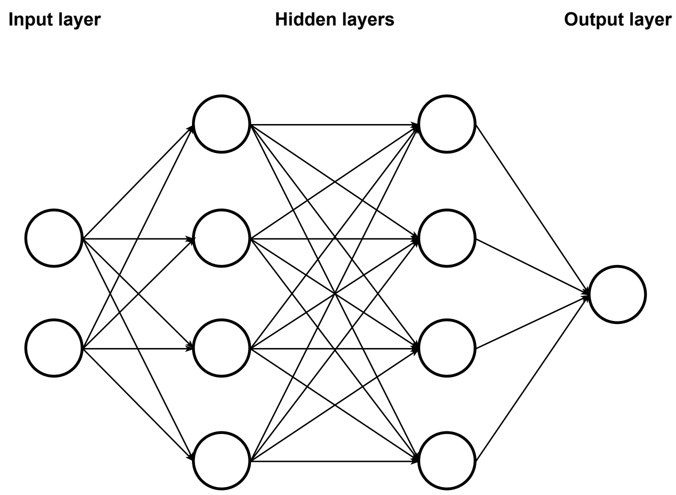

Figure 7. Perceptron network with hidden layers.

It can be theoretically demonstrated that any function can be computed using perceptron networks, no matter how complex they may be [317,318]. In this sense, perceptron networks are computationally universal, meaning that any computational operation can be represented through a network of perceptrons. Despite this strong theoretical guarantee, neural networks were not widely used in the scientific community for many years due to the lack of an efficient training algorithm. However, the landscape changed with the pioneering work of Rumelhart et al. [319], who introduced the sigmoid neuron and the backpropagation algorithm. It is worth noting that the backpropagation algorithm has been independently invented multiple times over the years in various domains, with its origins traceable back to the work of Werbos [320] in behavioral sciences [321].

The process of training a neural network using the delta rule involves adjusting its weights to minimize prediction errors, aiming to achieve the desired output within a minimal margin of error. For successful training, it is crucial that the relationship between the weights and the output is smooth, meaning that small changes in the weights should result in small changes in the output [317]. However, when using perceptrons as neurons, the binary step activation function (Equation (22)) introduces discontinuity, leading to an abrupt and discontinuous relationship between weights and outputs.

To address this issue, Rumelhart et al. [319] proposed the sigmoid neuron, which utilizes the sigmoid function as its activation function. It is worth noting that there is evidence suggesting the use of sigmoid activation functions in biological neurons [321]. The sigmoid function is defined as:

$$f(u) = \frac{1}{1 + \exp(-u)}. \tag{23}$$

Figure 8 demonstrates the comparison between the binary step and sigmoid activation functions. The sigmoid function can be interpreted as a smoothing of the binary step function, thereby facilitating the training process.

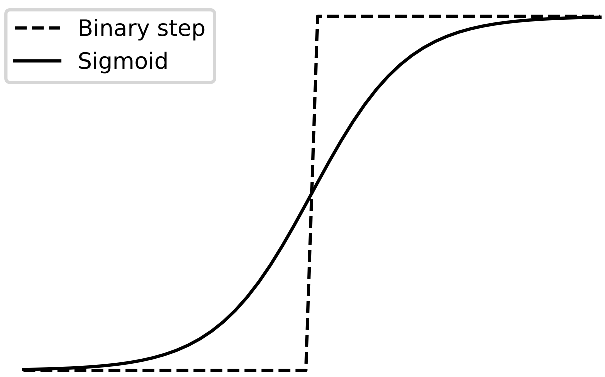

Figure 8. Comparison between the binary step and sigmoid activation functions.

Another major contribution from Rumelhart et al. [319] was the development of the error backpropagation algorithm. This algorithm consists of two stages in each iteration: a forward pass and a backward pass. During the forward pass, the network calculates its output, while during the backward pass, the contribution of each layer to the error of the preceding layer is determined (this process occurs in reverse, from output to input). The outcome of this backward pass is an efficient measurement of the error gradient throughout the network, enabling the application of the gradient descent algorithm to minimize errors.

Although the proposals made by Rumelhart et al. [319] generated significant interest in the scientific community regarding the utilization of artificial neural networks, challenges associated with training networks with multiple layers persisted for an extended period. It is only in recent times that these difficulties have been overcome through advancements in the field of deep learning. Deep learning is a methodology that uses multiple layers of information processing to represent complex concepts by associating simpler ones [298]. This approach is inspired by the way biological neural networks process information [321]. In the context of artificial neural networks, deep networks are characterized by having more than one hidden layer. Notable contributions in this field include those by Hinton et al. [322], who introduced a per-layer training strategy utilizing graphics processing units (GPUs), and Glorot and Bengio [323], who proposed alternative activation functions (in addition to the popular sigmoid) and new weight initialization strategies to address the problem of vanishing or exploding gradient across layers. Cutting-edge advancements in machine learning, such as the victory of the AlphaGo system in the complex Chinese board game Go [324,325] and the development of large language models like the generative pre-trained transformer (GPT) [326], predominantly rely on the application of deep neural networks.

One common criticism of methods in this class is the lack of interpretability of the results by humans, resulting from a limited understanding of the internal modeling mechanisms [285]. Recent studies [327,328] aim to address this concern through systematic investigations using carefully selected models. The issue of interpretability and its importance for future research is further discussed in Section 8.

For an extensive review of the field of deep learning, the reader is referred to Sengupta et al. [329]. Yu and Zhang [1] offer a review specifically focused on the application of deep learning in process monitoring. Zhou and Zhu [330] provide a critical review of the application of deep neural networks in modeling dynamical systems.

It is important to note that besides the realm of deep learning, other types of advancement have been proposed in traditional training algorithms. One notable example is the concept of extreme learning machines (ELM) introduced by Huang et al. [331]. ELM facilitates the rapid learning of neural networks by employing a random selection process for the nodes where the weights are trained.

4.1.2. Neural Network Architectures

A variety of architectures can be proposed for neural networks [315,332]. The most common distinction lies between feed-forward (or acyclic) networks and recurrent (or cyclic) networks. In feed-forward networks, information flows unidirectionally, from input to output, as illustrated in Figure 7. Conversely, recurrent networks process inputs sequentially, transforming them into hidden states, which are then retrofed back into the network. This enables the interaction of an input provided at time t with inputs provided at previous times $(t-1, t-2, \ldots)$. Consequently, recurrent networks prove particularly valuable for modeling time series, textual sentences, or biological sequences [315].

Attention mechanisms, proposed by Bahdanau et al. [333], aim to enhance the performance of deep neural networks by assigning weights to and combining input data based on their relevance to the given task. Particularly, self-attention, introduced by Cheng et al. [334], enables each position in the input data to independently relate to all other positions. This concept has revolutionized the field of deep learning, particularly in the context of sequential data, enabling the development of feed-forward architectures capable of modeling long-range dependencies using parallelizable training algorithms [335]. The Transformer architecture [336], widely employed in large language models such as GPT [326], is built upon self-attention mechanisms.

Convolutional networks are a type of neural network characterized by their sparse architecture, where filters are selectively applied to activate only the neurons essential for optimal learning [285]. This unique design has proven to be highly successful, particularly in computer vision applications, where convolutional networks have recently surpassed human capabilities [337]. By being inherently sparse, convolutional networks excel at learning and recognizing local patterns, which can be identified at various positions within different samples. In the context of image processing, these patterns can encompass edges, curves, and other geometric shapes, allowing convolutional networks to effectively capture the visual characteristics of images. This capacity to detect and extract meaningful features from images has made convolutional networks a cornerstone in various computer vision tasks, such as object recognition, image classification, and segmentation.

Autoencoders are neural networks specifically designed to generate, with certain constraints, output data that closely resembles input data [338]. The utility of this methodology results from the nature of these constraints. For instance, if the hidden layer contains fewer neurons than the input layer, the network is compelled to represent the data in a lower-dimensional latent space, referred to as the "code." This behavior is reminiscent of the latent variable modeling and dimensionality reduction process observed in the PCA model [339]. For a comprehensive examination of the application of autoencoders in process monitoring, please refer to Qian et al. [340].

Generative adversarial networks (GAN's), originally proposed by Goodfellow et al. [341], are composed of two neural networks: the generator and the discriminator. These networks are jointly trained to produce synthetic data that progressively resembles real data. The generator network aims to generate data that closely resembles real data and receives feedback from the discriminator network. The discriminator network, on the other hand, is trained to maximize its ability to distinguish between real and synthetic data. GANs have shown promise in various applications, including time series analysis and anomaly detection. Comprehensive reviews on the application of adversarial generative networks to these specific problem domains have been published [342,343].

Restricted Boltzmann machines utilize the Boltzmann learning rule along with topology constraints to enhance the efficiency of the training process [344]. A prominent class of deep learning networks that employ unsupervised learning is known as deep belief networks, which are constructed by composing restricted Boltzmann machines [345]. Moreover, restricted Boltzmann machines are commonly employed in the pre-training phase of conventional (supervised) deep networks [315].

Self-organizing maps combine the principles of competitive learning and Hebbian learning to create typically two-dimensional representations of complex patterns while preserving the topological structure of input data [346,347]. These representations facilitate pattern visualization, particularly in tasks such as clustering or latent variable modeling.

Fuzzy architectures, also known as neuro-fuzzy systems, employ neural networks as a learning mechanism to adapt parameters within a fuzzy logic system [15]. Fuzzy logic represents a generalization of Boolean logic, where the truth values of variables can take continuous values within the range of 0 (false) to 1 (true). It is important to note that fuzzy logic, independent of neural networks, can be utilized to design learning systems, as demonstrated by applications described in Section 4.2.

4.1.3. Applications in Process Monitoring

Applications of neural networks for monitoring purposes can be historically categorized into two approaches. The first approach, classified as supervised learning, involves assigning a scenario (normal or faulty) to each output node of the network. The network is then trained to classify a given dataset according to these scenarios [310,348,349]. In the second approach, classified as unsupervised learning, neural networks are integrated into traditional methods such as principal component analysis (PCA) to introduce nonlinearity [82,83]. Advancements in this second approach include the utilization of neural networks to extract nonlinear features independently from the formalism of traditional techniques, as proposed by Zhao and Lai [350].

There are works involving applications of process monitoring using several network architectures, such as recurrent networks [351–367], convolutional networks [299,349,368–382], autoencoders [282,365,379,383–405], generative adversarial networks [406–416], Transformers [417–425], restricted Boltzmann machines/deep-belief networks [345,412,426–432], self-organizing maps [433–441], neuro-fuzzy networks [442–451] and extreme learning machines [452,453], including recently in unsupervised form [454–456].

Besides specific works on process monitoring, ample perspectives on general applications of neural networks in chemical engineering have been published [332,457–459].

4.2. Other Techniques

The field of machine learning encompasses a rich body of literature, presenting numerous other techniques that deserve mention, including:

- **k-nearest neighbors** (**k-NN**): This method involves classifying a point in variable space based on the classes of its closest k neighbors. It is a simple and effective machine learning technique commonly used for datasets that are not excessively large and lack known probability distributions [15]. Several applications have utilized k-NN classification in the field of process monitoring, such as application to a semiconductor manufacturing process [460]; identification of faults that affect the entire plant through the application of the k-NN method for estimating delays [461]; diagnostic methods inspired by the ideas of reconstruction and contribution [462,463]; use of Dynamic time warping (DTW) algorithm [464–466] to define the Mahalanobis distance between the time series to be used in the k-NN technique [467]; and dealing with variable sample rates through sample grouping [468].
- **Support vector machines (SVMs)**: Support vector machines, originally proposed by Boser et al. [469], aim to identify hyperplanes that can effectively separate the variable space into distinct regions of interest [15,321]. These hyperplanes are determined based on a small subset of observations known as support vectors. To accomplish this, a concept called margin is defined, representing the distance between the hyperplane and the support vectors. The training process involves solving an optimization problem: in classification tasks, the goal is to maximize the margins, while in regression tasks, the objective is to minimize them. To introduce nonlinearity, SVMs utilize mappings through kernel functions [99] to transform the data into alternative mathematical spaces where hyperplanes can be employed to address the specific task at hand. Two methods have been proposed to expand the application of SVM to unsupervised classification scenarios: support vector domain description (SVDD) and one-class SVM (1C-SVM). SVDD, initially proposed by Tax and Duin [470], utilizes a hyperplane as the geometric form for separation, while 1C-SVM, introduced by Schölkopf et al. [471], employs a sphere. Recent applications of SVM in the field of process monitoring include: methodologies to select Gaussian kernel parameters for applications in fault detection [472]; integration of the SVM with Fisher discriminant analysis and wavelets for application in multiscale problems [236]; integration of the SVM with the recursive feature elimination technique for variable selection [473,474]; an adaptive methodology for application in non-stationary processes [475]; outlier treatment through a pruned SVDD algorithm [476]; a two-step SVDD method with genetic algorithm optimization to deal with dynamic, non-linear and non-Gaussian features [477]; development of an SVDD methodology to deal with incomplete data coming from industrial units with multiple local units [478]. It is worth noting that support vector machines (SVMs) have found broader applications in the field of mechanical engineering compared to chemical engineering [479,480]. This discrepancy may be attributed to the inherent characteristics of data generated by mechanical machines, which typically exhibit less redundancy and collinearity compared to data from chemical processes.
- **Tree-based methods**: The decision tree model constitutes a tree-like representation that encompasses various possible decision paths and their corresponding outputs [481]. The direct utilization of decision trees as machine learning models offers the significant advantage of complete transparency and interpretability regarding the decision-making process. Due to its simplicity, the application of decision trees in the field of process monitoring is relatively limited, with more prevalence in the realm of quality control [482,483]. A commonly employed strategy involves utilizing an ensemble of trees, such as in the random forest algorithm where multiple trees are trained using random subsets of the training data. The final result is determined by aggregating the values from each tree, often through averaging. This approach finds extensive use in process monitoring applications [484–495].

- **Manifold learning**: Within this category of techniques, the aim is to reduce the dimensionality of the data while preserving the underlying nonlinear relationships between its elements [496,497]. These methods operate by projecting the data onto manifolds, which are local topological subspaces that can be treated as approximately Euclidean. This mechanism proves valuable in capturing the local geometric structure of the data, surpassing the limitations of traditional latent variable models such as PCA, which only consider global Euclidean structures for dimensionality reduction. Specific models include t-SNE [498], which models similarities between data points using probability distributions, and UMAP [499], which uses a combination of graph-based and probabilistic approaches to achieve better global structure representation and computational efficiency. Examples of applications in the process monitoring field include: a method based on kernel functions specially designed for batch processes [500]; use of manifold learning in conjunction with clustering methods for temporal alignment and phase identification of batch processes [501]; a new monitoring index for control charts based on manifold learning [502]. A recent trend is the proposition of methodologies that simultaneously represent local and global structures in the data [503–509].
- **Sparse dictionaries**: Initially introduced by Olshausen and Field [510] for image recognition tasks, the concept of learning through sparse dictionaries involves the use of bases with dimensions greater than that of the original space, enabling the creation of a sparse representation of the data [511]. The basis itself is referred to as a dictionary, with each constituent element termed an atom. According to Mairal et al. [511], the utilization of sparse representation facilitates the capture of high-order correlations more effectively. Although the application of this methodology to monitoring problems is relatively recent, it holds promising potential. Examples of work in the field of process monitoring related to sparse dictionaries include: a sparse representation classification method for fault detection and diagnosis [512]; application of dictionaries to a semiconductor manufacturing process [513]; generation of sparse contribution graphs [514]; fault detection in multimodal processes [515]; use of kernel functions to capture nonlinearity [516]; a robust methodology for monitoring processes with multimodality, noise and outliers [517]; a methodology for decomposing large systems into modules, in order to facilitate the detection of small faults [518]; a distributed computing methodology to deal with big data [519]; an adaptive technique to deal with multimodal processes [520]; and a method for linear and nonlinear variable partition applied to dynamic process monitoring [521].
- **Bayesian networks**: Bayesian networks are graphical models represented by directed acyclic graphs (DAGs) that capture probabilistic dependencies among a set of variables [15,522]. This modeling framework excels in handling uncertainty and offers causal inference capabilities, enabling the description of relationships between causes and symptoms [523,524]. Examples of works in the field of process monitoring related to Bayesian networks include: use of specialist knowledge to build a Bayesian network for monitoring a tubular reactor [525]; construction and causal decomposition of the T^2 control chart [526]; monitoring of dynamic processes for propagation identification and root cause diagnosis [527]; use of Bayesian networks to filter the smearing-out effect that occurs in contribution methods for fault diagnosis [528]; development of a robust Bayesian network to deal with low-quality data [529]; development of a Bayesian classification network based on statistical alarms to deal with the effect of noise [530]; detection of kicks of hydrocarbons in drilling operations [531]; combination of mechanistic and state transition analysis to recognize propagation paths of process faults [532]; a graphical way to represent the causal interactions of a process using an interpretable Bayesian network [533]. A recent trend involves integrating Bayesian networks with traditional methods such as PCA and ICA [534–536]. For a tutorial review on applying Bayesian models to the process monitoring problem, refer to Raveendran et al. [537].

- **Methods based on fuzzy logic**: Fuzzy logic serves as an extension of Boolean logic, allowing for variable truth values to range continuously between 0 (representing false) and 1 (representing true). Monitoring systems that leverage fuzzy logic have been developed in various studies, including the proposal of an adaptive classification system capable of incorporating information about different modes of operation [538]; a fuzzy logic-based clustering algorithm applied to the Tennessee Eastman benchmark [539]; a combination of genetic algorithms and fuzzy logic-based clustering techniques to identify patterns in multivariate time series data [540]; the proposal of a novel clustering algorithm by integrating k-NN techniques with fuzzy k-means [541]. It is worth noting that fuzzy logic systems can be integrated with the learning capabilities of neural networks, as discussed in Section 4.1.2.

In addition to traditional methodologies, a recent trend has emerged in which techniques that are less limited by predefined model structures are employed. These approaches prioritize the exploration of intrinsic patterns inherent in the data itself. Several notable methods within this trend include:

- **Clustering**: In this approach, which is one of the most popular unsupervised learning methodologies, data is divided into clusters based on some similarity measure [542]. In the realm of process monitoring, various applications have emerged, including a *fuzzy* methodology to perform clustering based on wavelets and the PCA model [543]; a hybrid method that combines Euclidean and PCA-based similarity measures [544]; a new clustering algorithm combining the techniques **kNN** and fuzzy k-means [541].
- **Logical analysis of data** (**LAD**): The LAD technique, introduced by Hammer [545], involves the binarization of data followed by the discovery of interpretable patterns using boolean modeling approaches [546]. The key advantage of this technique lies in its ability to produce patterns that can be readily interpreted, facilitating the transformation of information into human understanding and knowledge. Applications of LAD in the field of monitoring have been recently proposed [362,547–551].
- **Empirical machine learning**: This approach revolves around the density of data in the variable space and yields algorithms that are entirely data-driven, devoid of probabilistic models and user-defined hyperparameters [552–556]. The theory introduces novel concepts such as cumulative proximity (a centrality measure), eccentricity (an anomaly measure), and typicality (a density measure). Applications in process monitoring have been developed [557–561].
- **Recurrence matrices and distance matrices**: Recurrence matrices and distance matrices are two-dimensional representations of data distance patterns. Applications of this methodology in process monitoring have been proposed in combination with various techniques, such as convolutional neural networks [371], recurrence quantification [562–565] and texture analysis [566–568]. Melo et al. [569] have recently proposed a novel methodology for visual process data analytics and process monitoring based on a direct analysis of distance matrix patterns.

5. Exploration, Characterization and Treatment of Process Data

In classical statistics, the theory is primarily developed based on the assumption that the data being analyzed originate from experiments specifically designed to test particular statistical hypotheses. On the other hand, historical process data arises from unplanned observations of real-world events, and the classical statistical literature strongly discourages the analysis of such data [570]. Consequently, applying classical techniques directly to historical data may not yield satisfactory results.

The field of statistics and data science that aims to explore and characterize data independent of any specific model is known as exploratory data analysis (EDA) [5]. In EDA, according to Morgenthaler, "one is free to choose any procedure to analyze the data, and the primary aims are to look at the data and to think about the data from many points of view". The application of EDA to process data has been discussed in the literature over the years. Pearson [570] described simple procedures, including quantile-quantile plots, data comparison plots, box plots, and moment characterizations, emphasizing the importance of such exploration. It was argued that unexpected data characteristics, such as outliers from different sources, can significantly invalidate results obtained through process monitoring techniques, particularly those based on quadratic error criteria. The author's main recommendations encompassed the following: (i) conducting adequacy tests for specific distributions, with particular emphasis on the Gaussian distribution, and comparing distributions across different datasets; (ii) employing robust statistical characterization methods, such as using the median instead of the mean, MAD (median absolute deviation from the median) instead of the standard deviation, and Spearman's correlation coefficient instead of Pearson's correlation coefficient; (iii) utilizing visualization methods such as box plots to compare variables. Abonyi [571] utilized box and quantile-quantile plots to analyze real historical data from a polyethylene production plant. The objective was to examine various manufacturing processes for a specific product and investigate the connections between different operational factors and the final product quality. Zhang et al. [260] proposed a data characteristics test to automate the selection of process monitoring techniques. Deviations from Gaussianity and linearity were used as indicators to aid the decision-making process regarding combinations and variations in the PCA and ICA models.

In addition to exploring process data, preliminary treatment procedures are crucial to ensure the effectiveness of modeling. In this context, Xu et al. [572] conducted an extensive review of the treatment of industrial data, referred to as data cleaning in their study. Four key steps in the data cleaning process were identified: filling in missing values, detecting outliers, removing noise, and performing time alignment/delay estimation. A wide range of methods available in the literature for each of these challenges was presented. The authors emphasized that the choice of data processing methods should be guided by the specific properties of the dataset itself, such as the proportion of missing data, as well as the existing knowledge about the process, such as the availability of a known model. Moreover, the methods employed should preserve and reveal the characteristics of regular data while effectively identifying the mechanisms that contribute to contamination, such as the nature of outliers.

In the context of robust methodologies, particularly concerning outliers and missing data, Zhu et al. [154] conducted a review from the perspective of big data. They examined commonly employed techniques for data preprocessing, as well as robust variants of PCA and other monitoring techniques.

An important step prior to applying a monitoring technique is the selection or generation of input variables. The area of machine learning that deals with this question is known as feature engineering. A notable approach gaining attention in recent years is statistical pattern analysis (SPA), proposed by Wand and He [573,574]. In SPA, statistical measures of process variables (mean, variance, autocorrelation, cross-correlation, among others) are selected as input variables for the modeling techniques (typically, PCA), which in many cases improves detection and diagnosis performance. Recent research on variable selection includes: the application of a criteria based on mutual information for variable selection [575]; an efficient technique based on sequential analysis of fault contributions [576]; a technique to extract non-stationary variables and its application to fouling detection [577]; the integration of variable selection methods with moving window techniques for multimodal monitoring, achieving good results on a particularly challenging fault in the TEP benchmark [578]; a comparison of shallow and deep learning methods for selecting variables in an industrial test platform focused on big data and Internet of Things concepts [579]; and the application of causality-based methods for feature selection in the TEP benchmark and in an oil platform fiscal metering plant [580]. Peres and Fogliatto [4] published a review on this subject, highlighting that the most used methods are LASSO regression, the genetic algorithm and regression by direct selection (forward selection).

Siang et al. [581] described good practices for acquiring and preparing process data. The authors emphasized the importance of (i) understanding the nature of the data through exploratory analysis; (ii) appropriately selecting the data subset to be used in the modeling; (iii) paying attention to open or closed loop conditions; (iv) align sampled signals at different rates; and (v) adjust the choice of metrics to the practical needs of the business problems.

Thibault et al. [582] conducted a review of data processing methodologies employed in process industries, focusing specifically on the pulp and paper sector. The authors divided these methodologies into three main categories: data cleaning, steady-state detection, and operating regime detection. Interviews with process specialists and software developers were conducted, leading to interesting insights regarding the gap between research advancements and practical implementations, among other points.

On the broad theme of exploration, characterization and treatment of process data, it is also worth highlighting the recent works by: Li et al. [583], which proposed dataset division based on the nature of the correlation (linear or non-linear) detected between variables and the application of a hierarchical strategy in these datasets to monitor linear and non-linear characteristics at different levels; Thomas et al. [584], who presented an integrated approach to data clustering and feature extraction, as an aid to increase knowledge about process data; Parente et al. [585], who proposed the application of the Monte Carlo technique to increase the amount of data used in training a monitoring model for a paper production process; Offermans et al. [586], which compared different methods for dynamically synchronizing process variables; and Rhyu et al. [587], which proposed an integrated methodology for automated outlier detection and missing data estimation.

6. Computational Tools And Benchmarks

6.1. Softwares

A wide range of computational tools exists for applying data-driven process monitoring techniques, with most of them being commercially developed. However, there is a lack of open-source options available. These options could significantly contribute to scientific progress, as exemplified by the remarkable advancements in the broader open ecosystem of data science and machine learning. This section specifically focuses on open-source tools for process monitoring within the field of PSE. Additionally, Appendix B provides a non-exhaustive list of some key commercial tools.

Open tools for PSE-based process monitoring are commonly developed as extensions within the Matlab environment. Camacho et al. [588] introduced MEDA, a specialized toolbox for multivariate analysis in the realm of chemometrics, featuring PCA and PLS methods, along with various monitoring applications. Jiang et al. [589,590] proposed the DB-KIT, designed for handling key performance indicators (KPIs) using regression algorithms and latent variable modeling. For batch processes, Gonzalez-Matínez et al. [591] introduced MVBatch, which offers notable functionalities such as data preprocessing, model cross-validation, and development on the GitHub platform. Yi et al. [592] presented Pre-Screen, a dedicated toolbox for pre-processing process data. Additionally, Villalba et al. [593] developed a graphical interface tutorial to facilitate comprehension of PCA-based monitoring. Schaeffer and Braatz [594] created a didactic Matlab application to aid in understanding latent variable models and regression. Melo et al. [5] proposed KydLIB, an open-source Python exploratory analysis package specially designed for historical process data. Rhyu et al. [587] made available in a GitHub repository a set of routines for automated outlier detection and missing data estimation. Furthermore, Sun et al. [595] and Alizadeh et al. [596] reported the existence of Matlab toolboxes for latent variable methods and causality analysis, respectively. However, the specific names of these tools and their download links were not provided in the referenced sources, and we were unable to locate them through internet searches.

As is apparent, there is a lack of open computational tools in the field of PSE industrial process monitoring, especially when compared to the wide range of options available in the data science environment. Moreover, the existing tools are mostly extensions of the Matlab environment, which hinders collaborative development due to its closed nature and restricts their usage in industries without the required software licenses. This issue is further discussed in Section 8.

6.2. Benchmarks

Benchmarks are standardized tests used to evaluate and compare the performance of computational systems, encompassing both hardware and software components [597]. In the field of PSE, a variety of open benchmarks exist to assess novel fault detection and diagnosis techniques. These benchmarks provide researchers with standardized datasets or simulation models that enable them to evaluate the efficacy and efficiency of their proposed methods. The remainder of the section provides a compilation of these benchmarks along with concise descriptions.

The Tennesse Eastman Process (TEP) is a chemical process simulation developed by Eastman Chemical Company to serve as a benchmark for evaluating process control and monitoring applications [49]. Originally, the TEP model was available as a Fortran-based process simulator, but it has now been implemented on different platforms. It continues to be the most widely used benchmark for process monitoring in the field of PSE [5].

PenSim is a simulation framework designed for studying the production process of penicillin. It incorporates an unstructured model of penicillin fermentation, initially introduced by Bajpai and Reub [598] and subsequently expanded by Birol et al. [599]. This particular process is the leading benchmark for evaluating methods developed specifically for batch processes [5].

DAMADICS (Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems) is a benchmark comprising both a process simulator and real datasets representing electro-pneumatic actuators used in a sugar production process [600]. The original simulator was developed using the MATLAB environment, while the real data was collected from the Lublin Sugar Factory in Poland. Due to its focus on a single equipment component, the DAMADICS benchmark is particularly suitable for evaluating model-based process monitoring techniques. However, there have also been studies that explore data-driven approaches within this benchmark [5].

The Benchmark Simulation Model No. 1 (BSM1) is a comprehensive specification of a simulation environment specifically created for evaluating control strategies in wastewater treatment plants [601]. Unlike the previous examples, BSM1 is a model specification, requiring users to implement it in their preferred programming language or simulation platform.

IndPenSim is a phenomenological model specifically designed to depict the fermentation process of Penicillium chrysogenum [602,603]. Unlike the conventional PenSim benchmark, IndPenSim stands out due to its industrial-scale representation, utilizing volumes of 100,000 L instead of the typical 100 L vessels employed in PenSim. Additionally, it benefits from the availability of real historical data and employs a structured model [604] that encompasses crucial aspects such as growth, morphology, metabolic production, and biomass degeneration.

The Cranfield Multiphase Flow Facility serves as a real-data benchmark for developing and validating statistical process monitoring procedures [129,605]. It represents a pilot-scale three-phase flow separation process, encompassing pipelines, a gas-liquid two-phase separator, and a water-oil-air three-phase separator.

The PRONTO (Process Network Optimization) heterogeneous benchmark dataset [606,607] builds upon the Cranfield benchmark, expanding its scope to address challenges associated with diverse data acquisition and storage formats. This dataset was specifically designed to incorporate a range of data types, including process measurements (such as high-frequency ultrasonic flow and pressure sensors), alarm records, events and change logs, an operation log, and video recordings.

The 3W dataset [608] offers a comprehensive collection of data encompassing real, simulated, and hand-drawn samples, which represent both normal and faulty operation periods of offshore oil wells across multiple well locations and operational timeframes. This benchmark dataset presents unique challenges inherent to real-world scenarios, including a substantial number of missing observations and frozen or incomplete variables.

BDsim is a biodiesel production process simulator that focuses on the continuous transesterification of oil with methanol [609]. It offers a comprehensive range of potential faults at different levels, including process, sensor, and actuator. Notably, BDsim surpasses existing simulators by providing previously unexplored scenarios like intermittent faults and equipment degradation.

The FCC Fractionator is a model designed to simulate a refining process that incorporates a fluidized bed catalytic cracker and a fractionator [610]. This recent proposal holds significant potential for fostering innovation in the field of process monitoring.

The Arc Loss Challenge is a benchmark that focuses on data from an industrial-scale mining and pyrometallurgy process [611]. This unique challenge introduces a comprehensive benchmarking workflow, featuring submission and evaluation pipelines in a competitive format.

For an extensive review and in-depth exploration of process monitoring benchmarks, the reader is referred to Melo et al. [5].

7. Industrial Applications

This section focuses on the application of data-based process monitoring in real industrial settings. It provides an overview of the historical development of industrial implementations in this field and describes the main works that have contributed to its advancement.

The application of the PCA technique in industrial settings dates back to the late 1980s. Wise et al. [11] introduced the use of PCA for process monitoring and demonstrated its efficacy through an application to data obtained from a ceramic smelter used for reprocessing waste in the nuclear industry. This study revealed that employing PCA on historical data could enhance the understanding of an industrial process, primarily by examining the contribution of each variable to the generation of principal components. Real-time industrial application results pertaining to this equipment were subsequently reported [612].

In the next years, numerous studies described the application of PCA and PLS in various industrial systems. Notably, Slama [613] investigated the operation of a catalytic cracking fluidized bed and the fractionation section of a refinery, shedding light on practical aspects associated with the analysis of real data. The study examined crucial factors such as the validity of data assumptions (e.g., absence of autocorrelation), sampling time, and its relationship with process time constants, among others. The findings of this work significantly contributed to advancing the understanding of the process. For instance, it revealed variations between different operating periods and identified a drift in the overall dynamics of the plant, despite the products being within the desired specifications. However, the presence of nonlinearities posed challenges to the performance of real-time monitoring systems.

During that period, a prominent area of application involved the utilization of multi-block and multi-way methods in batch and semi-batch polymerization systems, with many studies using real data [12,170,614]. However, these works primarily focused on the methodology and mathematical aspects of the techniques, rather than emphasizing their practical implementation. In contrast, notable contributions from the company DuPont stood out during the same period for their strong emphasis on case studies [63,615,616]. For instance, Piovoso and Kosanovich [616] addressed challenges related to the lack of representative data and the resistance of operating staff to conduct planned experiments that encompassed the entire operational range. This resistance primarily resulted from concerns about producing an excessive amount of products that fell outside the desired specifications.

In a survey conducted by Kourti and MacGregor [617], the applications of PCA and PLS methods up to that point were examined. The authors also introduced two novel applications specifically in continuous and semi-batch polymerization processes. The survey highlighted the rapid acceptance and adoption of multivariate statistical monitoring techniques within the industry. One of the case studies presented in the survey demonstrated how the application of these methodologies promptly identified variables that led to low purity and/or recovery. This task had previously consumed a significant amount of time and effort for the plant engineers. In another study, a fault resulting from a typing error made by an operator was successfully detected. However, it is important to note that real-time implementations of these methodologies were not presented in the case studies discussed.

In a highly influential article, Wise and Gallagher [13] provided a comprehensive and didactic presentation of the monitoring methodology using the chemometric approach to the process control community. This work served as a consolidation of the research conducted by the group in previous years. The article included several application examples to illustrate the methodology. These examples encompassed various domains, including (i) the ceramic melting equipment, which had been previously studied by the same authors [11,612]; (ii) monitoring the mitigation of nuclear waste in storage tanks, where subtle trends were identified through principal component analysis; (iii) monitoring a batch process utilizing near-infrared spectrometry (NIR), employing evolving factor analysis and multivariate curve resolution techniques. These application examples illustrated the effectiveness and versatility of the chemometric approach in process monitoring and control.

Following the influential work of Wise and Gallagher [13], there was a notable increase in the variety of research papers featuring examples of industrial applications using latent variables models. An example of this trend is demonstrated by Ignova et al. [618], who applied multi-way PCA and PLS to data from a fermentation bioprocess. Their study revealed that substantial cost reductions were possible by utilizing the developed models to determine feed quality and predict the concentration of the final product. Subsequent works reported the real-time application of the methodology, which was integrated into an expert system called the G2 real-time expert system [619,620].

Gurden et al. [621] reported the introduction of chemometric techniques in an industrial pilot plant. Several noteworthy results were achieved, including fault detection and diagnosis, the utilization of principal components to generate process trajectory graphs for monitoring plant operation, and the application of latent variables in cross-correlation analysis to enhance the interpretive capacity of analyzing time-delay effects. The authors highlighted the significance of their findings, as the gained experience and improved process understanding empowered plant personnel to propose the installation of new sensors and analyzers. This proactive approach was driven by the anticipated competitive advantage that these advancements would provide.

Taylor [622] documented the application of PCA for detecting instabilities in a blast furnace of British Steel. Notably, the implementation of PCA-based monitoring led to a significant increase in anticipation time compared to traditional detection methods. This work stands out for two interesting characteristics: firstly, it provides a detailed account of the real-time implementation within the plant, offering more comprehensive insights than previous studies; secondly, it is a notable industrial work conducted without the participation of university specialists as authors, marking a departure from the trend observed since the DuPont studies in the early 1990s.

Neogi and Schlags [623], from the Air Products and Chemicals company, applied MPCA and MPLS to analyze an emulsion polymerization batch process. A novel approach was adopted, where the trajectories of each batch were compared based on reaction extent rather than batch time, marking a significant departure from previous practices. Although real-time implementation was not carried out, the application of monitoring methods to historical data yielded successful fault detection and diagnosis. Additionally, a PLS model was developed, enabling the prediction of product viscosities. This enhanced process understanding and contributed to increased operational efficiency.

Martin et al. [624] presented an application of the PCA technique to data from a fluidized bed reactor and addressed crucial aspects that they believed were necessary for the widespread industrial adoption of these techniques. These aspects included: (i) developing monitoring schemes that require minimal amounts of process data; (ii) exploring the use of multi-block techniques to monitor the entire plant; (iii) developing generic models capable of monitoring multiple products or grades; (iv) investigating the impact of dynamics and control loops on the performance of monitoring systems and the potential for masking faults; (v) assessing the sensitivity of PCA and PLS and evaluating the suitability of their confidence limits; and (vi) extending the use of density-based confidence limits for monitoring dynamic processes.

Bissessur et al. [625] explored two approaches for fault detection in a papermaking plant. The first approach involved monitoring specific equipment using neural networks applied to vibration data. The second approach utilized PCA to monitor the overall process. Interestingly, the authors made an important observation that faults were better detected by analyzing the lower-order principal components. This finding challenged the conventional belief that higher-order principal components contain the most important information. In this case, the first components primarily captured variations in mass and energy balances as well as movements of setpoints in the plant. This insight highlighted the significance of considering the lower-order principal components for effective fault detection, contrary to the usual assumption about the relative importance of principal components in PCA analysis.

Dudzic et al. [626] described with unprecedented detail the implementation of PCA and PLS techniques in industrial processes at Dofasco, a steel production company. The implementation of these techniques has been ongoing since 1993, as described by the authors who were all employees of the company. The report focused on two specific applications: (i) continuous casting machine monitoring, and (ii) a control system for determining the optimal amount of reagents required for sulfur removal from pig iron. The authors emphasized the progression from offline analyses to online implementations throughout the 1990s, which was supported by close collaboration with the Advanced Control Consortium at McMaster University. The employed techniques offered a balance between ease of development and insights into the process, surpassing the limitations of purely phenomenological or empirical models. Additionally, these techniques demonstrated robustness against noise and missing data. The report highlighted the significance of data pre-treatment, including the selection and application of filters, grouping of measurements to eliminate trends, and handling of missing data. In the monitoring of the casting process, PCA was applied to reduce the dimensionality of the training matrix from 240 to 10 variables, considering the application of delays for dynamic treatment. The authors provided the operators with two control charts based on the T^2 index and one based on the Q index, implementing a progressive alarm scheme with yellow and red colors, respectively. The authors reported impressive results, including a 50% reduction in downtime after two years of operating the monitoring system and achieving the highest productivity in the machine's 12-year history.

Albert and Kinley [627] investigated the application of PCA and PLS techniques to monitor the biosynthesis of the antibiotic tylosin during batch fermentation. The authors encountered interesting challenges associated with poor data quality and variations in data quality across different locations within the plant. These issues posed significant obstacles for data pre-treatment and highlighted the difficulties in integrating monitoring technology into the plant since the variation in data quality added an extra layer of complexity to the monitoring process.

Karim et al. [628] documented their experience with the implementation of data-based methods in the bioprocess industry. The study shed light on specific challenges faced in this field, such as extracting biologically meaningful information from a limited number of real-time process measurements. One of the complications arises from the fact that the measured data primarily reflects the conditions inside the bioreactor, while the desired information pertains to the physiological state of the cells, which traditionally necessitates expensive offline analyses. The authors demonstrated how the utilization of MPCA helped alleviate these limitations by enhancing data interpretation capabilities.

Kumar et al. [629] reported the implementation of PCA in a high-pressure polymerization process. The study demonstrated that the technique effectively identified various multivariate faults, characterized by correlations, which were not detectable using univariate charts, where individual variables remained within their normal operating ranges. Notably, transitions in product grades and process changes were more clearly observed in the latent variable space. While the study mentioned the installation of a monitoring system at the plant, the specifics of the online operation were not described.

In 2003, a symposium titled "Abnormal situation detection and projection methods—industrial applications" took place in Ontario, Canada. Kourti [630] provided a comprehensive summary of the event, including descriptions of the presentations delivered. The symposium was attended by professionals from companies such as Dofasco, DuPont, Perceptive Engineering, and the COREM consortium (Mineral Research Consortium), as well as researchers from various universities. The following are brief summaries of the four lectures presented:

- Paul Nomikos from DuPont presented their successful experience with the utilization of PCA and PLS for offline analysis. They highlighted the effectiveness of these methods in tasks such as identifying and resolving sources of variability, enhancing phenomenological processes, and developing new control strategies. However, Nomikos also emphasized the challenges associated with implementing these techniques in an online setting. Specifically, the long-term maintenance of the models was identified as a significant hurdle that needed to be addressed to enable successful online implementation.
- Theodora Kourti provided practical recommendations for the application of data-driven monitoring systems. She stressed the significance of operator acceptance, noting that even within the same company, similar units may require different screens based on operator preferences. Kourti suggested starting implementations with small-scale problems rather than attempting to monitor the entire plant right from the beginning. Furthermore, she emphasized the importance of process knowledge, as it helps determine appropriate transformations and weights for variables, as well as assess the quality of obtained data. According to her findings, a linear model is typically sufficient for monitoring deviations. Kourti concluded her recommendations by identifying three motivating factors for industries to adopt multivariable monitoring systems: financial return, safety concerns, and compliance with environmental regulations.
- Barry Lenox emphasized that numerous monitoring projects represent pioneering endeavors within their respective industries. He noted that once the initial application is successfully implemented, the transfer of the monitoring system to similar processes becomes relatively straightforward. Lenox further shared his experience, indicating that the confidence of operators, engineers, and managers in these systems can grow to a point where automatic shutdowns or interventions can be programmed based solely on the information provided by the monitoring systems.
- Vit Vaculik from Dofasco highlighted the potential risks associated with inadequately developed monitoring applications. He discussed the various barriers that can hinder the development process, including challenges related to data availability and quality, the technical expertise of staff members, and the acceptance of new technologies within the organization. Vaculik also pointed out the sometimes challenging decision companies face when deciding between purchasing pre-existing monitoring systems or developing their own solutions in-house.

Miletic et al. [631] from Dofasco and Tembec presented an industrial perspective on the online implementation of PCA and PLS techniques. The authors proposed a comprehensive method for developing monitoring systems, consisting of offline and online steps. The offline step involves data selection and model preparation, development, and evaluation. Despite being potentially overlooked by academic researchers, the authors emphasized that data selection is an expensive and vital process for the success of the application. The online stage focuses on system development, integration, and maintenance. The authors highlighted the importance of close collaboration with operators during this phase to understand their requirements and ensure a smooth interaction with the monitoring system. To illustrate the proposed methodology, the authors provided three application examples. First, they demonstrated the use of PCA for fault detection in a continuous casting machine. Second, they applied PLS to develop predictive models for a desulphurization process. Finally, they employed PCA for fault detection in a continuous digester. These practical examples emphasized the importance of a systematic approach to developing and implementing monitoring systems in real-world applications.

Kourti [632] presented a review of industrial monitoring applications utilizing multivariate statistical models. Interesting topics covered include dealing with future observations of missing data, adaptive models, monitoring batch processes, monitoring transitions, control using multivariate statistical models, image analysis and archiving, and data compression and reconstruction.

In the work conducted by Zhang and Dudzic [633], another implementation project at Dofasco was described. This particular implementation introduced a novel trajectory synchronization scheme for monitoring transition operations. Similar to previous publications from the company, the study provided insightful details regarding the implementation process, including the operation of alarms and control system screens.

Qin et al. [634] presented a hierarchical and comprehensive framework for the control and monitoring of new-generation semiconductor factories at that time. One of the notable findings highlighted in the study was the relatively greater significance of drifts compared to sensor faults in semiconductor processes.

Chiang and Colegrove [635] presented a comprehensive application of robust PCA within the industrial setting of Dow Chemical. The focus of their investigation was to effectively monitor the production process of solid epoxy resins, a specific product manufactured by the company. The authors provided valuable insights into practical considerations, including model performance, maintenance and transfer. Notably, they detailed the implementation of an online monitoring system, which soon after its deployment proved instrumental in detecting quality issues with the raw materials.

Kano and Nakagawa [636] presented significant advancements and practical implementations of data-based monitoring in the steel industry, drawing on examples from the operations of Sumitomo Metals Kokura. The researchers proposed a novel approach that effectively handled qualitative data, while also establishing connections between operational conditions and production yield. Notably, their methodology enabled the simultaneous analysis of multiple processing units, thereby offering a comprehensive perspective on the manufacturing process.

Miletic et al. [637] provided further insights and experiences about online applications at Dofasco. The authors specifically highlighted the development of a flexible computational framework that facilitated data manipulation and integration of new models. This framework enabled seamless incorporation of various techniques beyond PCA and PLS, such as Bayesian networks and image analysis techniques.

AlGhazzawi and Lennox [638] presented a study focusing on the development of a monitoring system for an oil refining process at Saudi Aramco. The primary objective of the study was to effectively monitor the non-stationary dynamics of the process. To achieve this, the researchers employed a recursive PCA method and also a multiblock approach for improved fault isolation. Notably, the authors emphasized the significance of incorporating the domain knowledge of plant personnel to ensure proper development and implementation of the monitoring system. It is worth mentioning that a real-time implementation of the system was not reported in this study.

Vanhatalo [639] provided a description of the application of the PCA technique for monitoring a pilot-scale blast furnace. The engineers overseeing the project acknowledged several advantages of this approach. Firstly, the application of PCA provided a comprehensive overview of the equipment's thermal state, thereby enabling a standardized evaluation of its condition. Additionally, it facilitated the establishment of formal criteria for determining the appropriate timing of control actions. However, the engineers also highlighted a few drawbacks. They mentioned the high costs associated with manpower required for model development, maintenance, and interpretation. Moreover, they noted that the multivariate representations used in PCA introduced unnecessary levels of abstraction, potentially making it more challenging to interpret the results accurately.

Darkow et al. [640] detailed the utilization of the PCA method for early detection of blockages or obstructions in an industrial stripper. The authors identified several key factors contributing to the success of the application. Firstly, they emphasized the importance of implementing preprocessing steps prior to model generation, such as variable selection based on expert knowledge and removal of univariate outliers through visual inspection. Secondly, they highlighted the significance of involving both plant personnel and information technology experts from the project's inception, ensuring a collaborative approach that incorporated domain expertise and technical knowledge. Lastly, the researchers underscored the development of a user-friendly graphical interface that catered to the needs of different users.

Dumarey et al. [641] presented an application of the PCA method for real-time monitoring of the continuous synthesis of an active pharmaceutical ingredient. The authors addressed the challenges surrounding the adoption of real-time monitoring methods within the pharmaceutical industry, highlighting the hurdles associated with integrating such methods into existing rigid quality control systems. Despite these challenges, the proposed application of PCA yielded significant benefits. It led to an enhanced understanding of the synthesis process, including a deeper comprehension of natural variability. Additionally, the monitoring system facilitated the identification of different fault types and their underlying causes.

Patwardhan et al. [642] discussed various applications of data science within the context of Saudi Aramco. Notably, some applications were related to process monitoring, including the assessment of controller performance and predictive maintenance of equipment. The authors emphasized the significance of utilizing data analysis techniques in effectively managing alarms, thereby enhancing the overall efficacy of monitoring systems when working in conjunction with operators.

Klanderman et al. [643] presented a case involving the application of the PCA technique to develop an adaptive monitoring methodology specifically designed for effluent treatment plants. The authors discussed several practical aspects related to this application. Firstly, they highlighted the enhanced understanding of process changes achieved by operators. However, it was noted that there could be instances where the results obtained from the monitoring system contradicted the operators' assessments, even when the methods were functioning as expected. As a recommendation, the authors suggested that operators and specialists should review the training data beforehand. Additionally, it was advised to reset the monitoring system after detecting a fault in order to prevent contamination with non-normal data.

Kumar et al. [644] described the development and application of a replicable system that utilizes multivariate statistical models for monitoring reformers in hydrogen production plants. The authors emphasized the significance of regularly retraining the models to ensure their effectiveness. To achieve this, they proposed a methodology that incorporates user feedback regarding false positives.

Based on the descriptions provided in the preceding paragraphs, it is clear that the main publications documenting practical applications of monitoring techniques were published during the 2000s. Subsequently, there has been a significant increase in studies on process monitoring from the 2010s onwards, with a growing emphasis on the utilization of real-world data. However, most of these studies primarily focus on offline analysis of new techniques or variations in traditional techniques applied to such data, while paying limited attention to practical implementation aspects. Furthermore, despite the increasing use of real data, simulated data, particularly from the TEP benchmark, remains prevalent in most research works. It is also worth noting that the main articles describing practical implementations predominantly rely on multivariate statistical models, particularly PCA and PLS, indicating a limited adoption of alternative techniques in this field.

Regarding recent studies involving offline analysis of real historical data, there are some noteworthy works that deserve highlighting, such as: the development of a methodology for detecting the decomposition of reagents in a copolymerization autoclave using PCA [645]; the application of machine learning methods for decentralized monitoring of an effluent treatment plant [646]; the application of PCA to monitor the physicochemical properties of crude oil blends [647]; the application of CCA to historical process data for detection and diagnosis of a sensor fault in a fiscal metering station of an oil processing plant, highlighting the potential for economic gains which may result in such a real-time application [648]; the application of different latent variable models to develop a strategy aimed at an increasing understanding of biorefineries processes [649]; the extension of recursive PCA with big data methodologies for monitoring a fluorochemical plant [650]; the development of a virtual sensor based on machine learning to predict faults in a metallurgical process [651]; the development of a system that combines various techniques such as recurrent neural networks and PCA to predict abnormal conditions in catalytic cracking processes [652]; the application of neural network-based nonlinear PCA for fault detection in gas turbines [653]; and the application of autoencoders for unsupervised monitoring of blast furnaces [405].

8. Perspectives and Challenges

The main perspectives and challenges in process monitoring research for the coming years are described in this section. These challenges can be broadly categorized into three key areas: system performance, significance/usefulness, and development environment. We believe that these categories encompass the primary obstacles encountered by researchers and practitioners in the field of process monitoring. Figure 9 shows a schematic illustration of the perspectives described along the section.

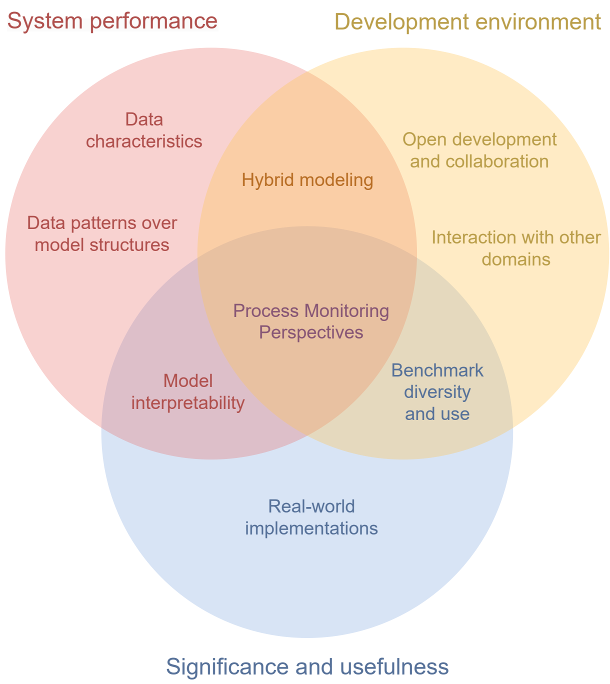

Figure 9. Schematic illustration of the perspectives for process monitoring divided into three overlapping categories.

8.1. Data Characteristics

The process monitoring research community primarily focuses on addressing challenges posed by specific data characteristics. Many studies have presented inventive approaches or modifications to existing methods, with a focus on specific data traits like dynamic behavior, nonlinearity, and non-Gaussianity. Several of these studies have already been discussed in the preceding sections of the present article. However, while it is important to introduce new techniques, we feel that the research community often overemphasizes this aspect, potentially overshadowing other critical challenges that also demand attention. The following subsections highlight these additional challenges.

8.2. Hybrid Modeling

The fusion of diverse data-driven methods, or the integration with other process monitoring approaches such as knowledge-based or phenomenological methods, has been widely discussed as a promising future direction in various review and perspective articles [15,165,654–661]. Md Nor et al. [15] conducted a comprehensive review of hybrid methodologies, emphasizing the necessity for such approaches due to the inability of individual methods to handle all potential faults that a process may encounter. Bi et al. [654] emphasized the challenges associated with integrating different methods, particularly in terms of leveraging the strengths of each methodology and establishing a robust decision-making process. Severson et al. [165] assert that hybrid approaches hold the most promise for addressing complex scenarios within the context of process monitoring. Sansana et al. [660] drew interesting conclusions from their literature review, for example, identifying that data-driven models have been used to enhance phenomenological models, but the reverse path of using phenomenological models to improve data-driven models is not yet fully explored. Bradley et al. [659] conducted an extensive review focusing on methodologies that integrate data-driven and phenomenological approaches, categorizing them into three key areas: hybrid modeling, physics-informed machine learning, and model calibration. Hybrid modeling involves constructing models that combine different phenomenological and data-driven submodels. Physics-informed machine learning entails incorporating physical knowledge into the structure of data-driven models. Model calibration refers to using data-driven models to adjust computer simulations and align them with experimental data. The authors provide a comprehensive comparison of these methodologies, emphasizing their respective advantages and limitations.

In addition to enhancing performance, the utilization of mechanistic models holds significant importance in improving model interpretability, a topic that is further discussed in the next subsection.

8.3. Model Interpretability

The importance of model interpretability as a research subject for the future has been recognized in various surveys [16,654,662–665]. Yan et al. [662] specifically highlighted the challenge of developing a generalized methodology to incorporate interpretability into deep learning methods. They emphasized the significance of achieving this objective to create optimal diagnostic models while acknowledging the trade-off between model interpretability and performance. Bi et al. [654] acknowledged the need for incorporating both human and machine intelligence to develop smart systems, emphasizing the role of interpretability to make humans trust in machine intelligence. The authors identified three approaches to achieve interpretability: intrinsic interpretable models, such as Bayesian neural networks [533]; enhancing black-box models, like deep neural networks, with interpretability capabilities [327,328,666–668]; and employing model-agnostic methods [390,669,670] such as LIME (local interpretable model-agnostic explanations) [671] and SHAP (Shapley additive explanations) [672]. Several recent publications have surveyed interpretability techniques for machine learning models [673–679]. In particular, Carter et al. [679] specifically examined applications in the process industries, focusing on risk management and standardization issues.

Among recent research on model interpretability, it is worth highlighting the works of: Sivaram et al. [327] and Das et al. [328], who investigated hidden representations in neural networks applied to classification and regression tasks, respectively; Harinarayan and Shalinie [680], who utilized the SHAP method to explain the results of a XGBoost [681] model applied to the TEP benchmark; Yang et al. [533], who developed an unsupervised Bayesian network model for monitoring processes at both global and local levels; Bhakte et al. [682], who proposed a methodology using alarm limits for explaining process monitoring results from deep learning models; Ye et al. [683], who combined the use of frequency spectra inputs with a layer-wise relevance propagation strategy to explain predictions of convolutional neural networks; and Melo et al. [569], who proposed a visual and interpretable methodology based on distance matrix patterns.

8.4. Data Patterns over Model Structures

New models that prioritize intrinsic patterns in data rather than specific mathematical structures and hypotheses are increasingly being utilized in process monitoring research, as discussed in Section 3.4. This approach has been extensively discussed by Angelov and Gu [556], who highlighted the shift from an assumptions-centered approach to a data-centered approach. According to the authors, this transition offers several advantages, including enhanced autonomy in the modeling process, adaptability to dynamic changes, and improved interpretability. We believe that the process monitoring community would greatly benefit from more research conducted with this mindset, as it can help address the challenges that arise when model assumptions do not align with the specific characteristics of process data.

8.5. Real-World Implementations

The process monitoring literature is dominated by articles focusing on proposing new mathematical models, while the scarcity of studies reporting real-world implementations is evident. Additionally, as discussed in Section 7, most of these implementations revolve around multivariate statistical models, particularly the PCA model, which suggests the existence of unexplored opportunities concerning the practical application of more recently developed techniques such as deep neural networks.

Unfortunately, many researchers lack access to real industrial facilities that enable the study of practical implementations, thereby aggravating the challenge of bridging the gap between industries and academia. In this context, Bi et al. [654] highlights the criticality of high-quality data availability over model complexity, despite the research community's predominant emphasis on the latter. To mitigate this issue, one potential solution could be to use the real-world dataset benchmarks proposed in recent years more frequently [5]. This matter is further discussed in the next subsection.

8.6. Benchmark Diversity and Use

In recent years, several new benchmarks for process monitoring have been proposed, as discussed in Section 6.2. However, despite these alternatives, the research community continues to predominantly rely on the well-established TEP benchmark [5]. This raises an important question: why does the TEP benchmark remain the preferred choice, even though the underlying problem it represents has already been addressed by solutions proposed over more than two decades of research?

One potential explanation for the continued popularity of TEP lies in its familiarity and ease of use. Researchers and practitioners have become accustomed to working with this benchmark, making it a comfortable and convenient choice for assessing new models and techniques. Furthermore, the extensive documentation and well-established methodologies contribute to its widespread adoption. However, it is important to critically evaluate whether this reliance on TEP hinders progress in the field. For instance, a significant portion of the literature primarily focuses on achieving marginal performance improvements in fault detection and diagnosis metrics. The proliferation of new models that offer only incremental enhancements may not truly benefit the research community and industry. Moreover, it is important to acknowledge that the TEP benchmark is based on simulated data, which presents significantly fewer challenges compared to real-world data.

Therefore, it may be more beneficial to explore alternatives that introduce challenges beyond those extensively investigated with the TEP benchmark. By shifting attention to new benchmarks, the research community can study different aspects of process monitoring, thereby expanding the field's knowledge in a more comprehensive manner. The alternative benchmarks encompass various data characteristics, such as nonlinearity, dynamics, and non-Gaussianity, which tend to be more pronounced compared to TEP. Additionally, real-world data substantially differ from simulated data, particularly in terms of data quality and variability. These are important aspects that should be addressed in future research through the utilization and development of more diverse benchmarks. For a more thorough discussion, the reader is referred to Melo et al. [5].

8.7. Open Development and Collaboration

Open development greatly accelerates scientific progress by promoting collaboration among researchers, facilitating synergistic contributions, and avoiding redundant work, among other advantages. In the field of PSE, a particular benefit is the potential to bridge the gap between academia and industry. For example, when implementing process monitoring procedures, companies heavily rely on commercial software solutions that often have limited connections to the cutting-edge methodologies developed in academia [582].

By embracing open development practices, universities and research centers can create open-source tools accessible to both academics and industry professionals. This can lead to the utilization of more advanced techniques in industrial applications, driving innovation, and enhancing the practicality of academic research. This phenomenon is already observed in domains such as data science and machine learning, where major corporations like Google and Facebook actively promote the development of cutting-edge and widely used tools such as Keras [684] and PyTorch [685].

An example of institutional support for open development in the process industries is Petrobras' Connections for Innovation program. This program aims to integrate the company with the open innovation ecosystem. One of this program's components, the Open Lab Module, is particularly relevant in the present context as it promotes collaboration to address challenges in the oil, gas, and energy sector through open-source projects. The 3W Project, mentioned in Section 6.2, is an example of an initiative within the Open Lab Module.

Specifically in the domain of PSE, Matlab has historically been the dominant programming language for developing innovative methodologies [686]. However, Matlab's proprietary nature, closed-source code, and licensing requirements make it less suitable for the open innovation paradigm. Therefore, we believe that transitioning to open-source tools like Python would be advantageous for the PSE community. This shift would not only facilitate the adoption of research findings by the industrial sector but also promote cross-domain integration, a topic discussed in the following subsection.

8.8. Cross-Domain Integration

In the field of PSE, the term process monitoring has been often used interchangeably with the term fault detection and diagnosis [20]. However, in other domains, such as signal processing and medical data analysis, the same concept is known by different names, such as anomaly detection [687–709], novelty detection [710–724] and change-point detection [725–741]. Although these three terms refer to variations in the same fundamental problem (at least from a computational and mathematical perspective), there is a noticeable lack of information exchange between the communities focused on these studies and the process system engineering community. This subsection highlights the potential opportunities for fostering interaction between these groups.

The differences in nomenclature contribute, at least partially, to the limited communication aforementioned. Moreover, the relative isolation can be attributed to the nature of the analyzed data and the methodologies utilized to address these problems. For instance, in the field of disease diagnosis from medical data, methods are typically developed for batch analyses [729], since the detection of a disease (anomaly) most often does not require the real-time and online analysis of a sequence while it is being generated. Moreover, medical data is often univariate, involving variables such as heartbeats [736]. These two conditions are not typical of PSE problems. Furthermore, process data exhibit specific characteristics, in particular high redundancy and collinearity between variables [24].

One could argue that this premise holds partially true, as the nature of problems in different domains has not undergone significant changes. However, with the disruptive phase that all domains are undergoing, particularly the industrial sector with the advent of Industry 4.0, the almost complete segregation of their literature on the same topic is becoming increasingly questionable. It is highly improbable that no potential for a productive exchange of ideas exists, even if the specifics of solution methods differ. The argument is also becoming increasingly unrealistic: works such as the one conducted by Alippi et al. [726], which examines the theoretical limitations of change detection methods applied to continuously generated multidimensional data, clearly share scope with PSE problems.

It is worth noting that other fields in industry and engineering employ similar terminology to address the monitoring problem, such as the term "fault detection and diagnosis". While communication between the PSE domain and these fields is generally better, we believe that it remains unsatisfactory. Notable examples of such fields include mechanical engineering [742–744], nuclear engineering [745,746], electrical engineering [747,748], and mechatronics [749,750].

In the realm of data science, the analysis of continuously generated data is commonly referred to as streaming data [751–753]. In this context, "concept drift" is a term used by data scientists to describe the evolution over time of patterns and relationships captured by machine learning techniques [754–759]. This idea could be of significant value for PSE, as process drift constitutes one of the most important process states and can usually be associated with process malfunctions [760]. This nomenclature has recently been adopted by research in the PSE field [761,762]. A current trend is the utilization of ensembles of methods to enhance efficiency in adapting to data evolution [763–766]. The streaming data literature offers potentially valuable works on clustering [767], pre-processing [768], outlier treatment [769,770], and event prediction [771], although we were unable to identify mutual references in the analyzed literature.

Causality analysis is an example of an application based on a time series that has gained prominence in the PSE field, as evidenced by numerous studies [580,772–782]. This technique uses statistical tests, such as the Granger causality test, to determine whether a given time series is useful in predicting another. Despite the significant advances made in the PSE literature, there seems to be a lack of attention from other fields towards causality-related works in PSE. This is particularly evident in recent reviews on causality conducted by Shojaie and Fox [783] and Guo et al. [784], as they do not reference any recent PSE studies.

Nonlinear dynamic analysis, with its rich history in both PSE [785–790] and time series analysis [791–794], holds immense potential for further applications in the evolving era of Industry 4.0. For instance, Aldrich [81] describes methods based on phase spaces, such as correlation dimension, Lyapunov exponents, and information entropy, emphasizing that adoption of these methods in industry is still limited due to the impact of noise on the quality of the descriptions. However, the fact is that in the domain of process monitoring, in particular, the potential use of nonlinear tools for the representation and evaluation of process states remains largely unexplored. As a matter of fact, the nonlinear characteristics of a time series can provide information about the process health, indicate possible process changes and suggest the use of different numerical tools for qualitative and quantitative description of the process behavior [81].

9. Conclusions

This paper presented a comprehensive overview of research on data-driven process monitoring in the PSE literature. Process monitoring has become increasingly important in the context of digital transformation, attracting attention from both academia and industry. It was observed that a large part of academic research is dedicated to developing data-driven models that can handle specific characteristics of process data, such as dynamic and nonlinear behavior. Furthermore, there is a recent trend toward exploring innovative machine learning methodologies. However, industrial applications still rely mainly on conventional statistical techniques, especially the PCA model. This reveals a gap between academic and industrial practices, but also an opportunity for transferring novel concepts and techniques from academia to industry. Furthermore, additional challenges and directions for future research were identified, such as assessing the relevance and applicability of academic proposals in real-world settings and fostering a collaborative and supportive environment among researchers and practitioners that enables the adoption and implementation of research results in practice.

Author Contributions: Conceptualization, A.M., M.M.C. and J.C.P.; methodology, A.M., M.M.C. and J.C.P.; investigation, A.M.; data curation, A.M. and M.M.C.; writing—original draft preparation, A.M.; writing—review and editing, A.M., M.M.C. and J.C.P.; supervision, M.M.C. and J.C.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: Afrânio Melo was employed by Petrobras, and Maurício Melo Câmara was employed by OPTIMATECH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. List of Surveys and Books on Data-Driven Process Monitoring

Table A1. Review and overview articles about process monitoring in PSE literature, in chronological order.

Reference	Description
MacGregor and Kourti (1995) [12]	Distinguishes quality control (multivariate extensions of Shewhart charts) from process control (latent variable models, PCA, and PLS).
Wise and Gallagher (1996) [13]	Explores multivariate statistical models (PCA and PLS) and provides an in-depth analysis of regression techniques, including less commonly cited methods such as rigid regression.
Dash and Venkatasubramanian (2000) [795]	Reviews the challenges associated with industrial applications of process monitoring techniques and provides an overview of the methods then available to address these challenges.
Kourti (2002) [24]	Examines multivariate statistical models (PCA and PLS) and covers practical aspects including database characteristics, sampling frequency, and other relevant considerations.
Qin (2003) [14]	Explores multivariate statistical models, specifically PCA, with a focus on comparing contribution and reconstruction methods, as well as investigating detectability, reconstructibility, and identifiability in various studies.

Table A1. *Cont.*

Reference	Description
Venkatasubramanian et al. (2003) [313]	Examines multivariate statistical models (PCA and PLS) along with qualitative trend analysis, statistical classifiers, and neural networks. Compares these techniques with model-based methods (observers, Kalman filters, parity, and qualitative model-based methods) discussed in the previous parts of the article series.
Miletic et al. (2004) [631]	Explores multivariate statistical models (PCA and PLS) with a specific focus on their industrial applications. Additionally, presents a methodology for developing effective and sustainable industrial monitoring systems.
Ganesan et al. (2004) [796]	Reviews strategies for process monitoring utilizing wavelets.
Kourti (2005) [632]	Explores multivariate statistical models (PCA and PLS) and addresses specific challenges encountered in real-world applications, including in-line calculations, control and archiving applications, as well as data compression and reconstruction.
Qin (2009) [25]	Its scope is similar to the previous work by the same author [14].
Yao and gao (2009) [68]	Reviews strategies for batch and multistage process monitoring.
Kadlec et al. (2009) [760]	Explores techniques and applications based on data-driven soft sensors.
Kadlec et al. (2011) [797]	Examines adaptive mechanisms for building data-driven soft sensors.
Ding et al. (2011) [798]	Examines multivariate statistical models (PCA and PLS) and provides a comparison with model-based methods such as Kalman filter, observers, and parity.
Sliskovic et al. (2011) [799]	In the context of models for application in soft sensors, it categorizes the methods into multivariate statistical methods, artificial neural networks, support vector machines, hybrid methods, and adaptive methods.
Das et al. (2012) [10]	Classifies data-driven methods into two categories: statistics-based (including univariate methods and multivariate latent variable methods) and artificial intelligence-based (including neural networks and fuzzy logic). Compares these techniques with methods based on a priori knowledge such as parameter estimation, observers, and parity.
Qin (2012) [18]	Its scope is similar to the previous works by the same author [14,25], with the addition of the PLS model.
MacGregor and Cinar (2012) [26]	Offers an overview of the application of multivariate statistical methods in the domains of monitoring, fault-tolerant control, and optimization.
Ge et al. (2013) [800]	Provides a process-based classification of methods, categorizing them into continuous, batch, non-Gaussian, nonlinear, time-variant/multimodal, and dynamic processes. Offers an extensive collection of methodologies not commonly cited in previous reviews.
Dai and Gao (2013) [658]	Reviews fault detection and diagnosis methods with a focus on data processing approaches.
Khatibisepehr et al. (2013) [801]	Reviews the design of soft sensors utilizing bayesian methods.
Yin et al. (2014) [28]	Examines multivariate statistical models (mainly PCA and PLS, although other models are mentioned) with a specific emphasis on modifications induced by industrial operating conditions, particularly addressing issues of uncertainty and dynamic behavior.
Yin et al. (2015) [802]	Focuses on multivariate statistical models and the integration with model-based techniques.
Severson et al. (2016) [165]	Focuses on multivariate statistical models, with a brief mention of other approaches, notably SVM.
Tidriri et al. (2016) [803]	Provides a comparative analysis between data-driven and model-based methods, with a specific emphasis on exploring hybrid methodologies. Data-driven methods are categorized into Bayesian networks, neural networks, univariate control charts, and multivariate statistical methods (specifically PCA and PLS).
Weese et al. (2016) [17]	Presents a review of machine learning methods, encompassing multivariate statistical methods, for the generation of control charts applied across diverse domains.
Shu et al. (2016) [804]	Explores challenges and perspectives in managing abnormal situations from the perspective of big data.
Souza et al. (2016) [805]	Reviews soft sensor methods for application in regression problems.

Table A1. *Cont.*

Reference	Description
Ge (2017) [130]	Reviews strategies adopted for monitoring plant-wide processes, focusing on multivariate statistical methods.
Reis and Gins (2017) [806]	Explores perspectives and historical trends in the application of process monitoring techniques, spanning from detection to diagnosis and prognosis.
Alauddin et al. (2018) [807]	Bibliometric analysis of the data-driven process monitoring field, with a focus on multivariate statistical methods.
Wang et al. (2018) [808]	Presents a systematic review of the application of multivariate statistics in process monitoring from 2008 to 2017, with a specific focus on Chinese literature.
Peres and Fogliatto (2018) [4]	Reviews strategies employed for variable selection.
He and Wang (2018) [809]	Presents perspectives and insights for process monitoring in the context of big data.
Quiñones-Grueiro et al. (2019) [103]	Reviews strategies adopted for multimodal processes.
Aldrich (2019) [81]	Reviews strategies adopted for dynamic processes.
Jiang et al. (2019) [131]	Reviews strategies adopted for distributed monitoring of plant-wide processes.
Reis (2019) [123]	Reviews strategies adopted for multiscale process monitoring.
Rendall et al. (2019) [67]	Reviews strategies for batch process monitoring, mapping them to a scale of modeling and implementation complexity.
Md Nor et al. (2019) [15]	Classifies the approaches into multivariate statistical analysis and machine learning, emphasizing the distinctions in method applications based on the type of fault being analyzed.
Pilario et al. (2019) [101]	Provides a review of the applications of the kernel technique for feature extraction in process monitoring.
Diez-Olivan et al. (2019) [657]	Reviews applications of data fusion and machine learning techniques in the context of prognostic activities.
Gopaluni et al. (2020) [810]	Reviews the applications of machine learning techniques for process monitoring and control.
Apsemidis et al. (2020) [99]	Provides a review of the applications of the kernel technique in process monitoring.
Park et al. (2020) [656]	Categorizes methods based on the type of system they are applied to, including dynamic, nonlinear, non-Gaussian, multimodal/time variant, and non-stationary systems.
Qin et al. (2020) [665]	Reviews multivariate statistical methods in the context of dynamic processes and introduces a novel mathematical framework aimed at unifying existing approaches.
Cohen and Atoui (2020) [811]	Provides a review of strategies for process monitoring utilizing wavelet analysis.
Ahmed et al. (2020) [812]	Conducts a bibliometric analysis focusing on techniques classified as based on Artificial Intelligence, excluding multivariate statistical methods such as PCA and PLS/CCA according to the authors' criteria.
Jiang et al. (2020) [664]	Presents a review of soft sensors, with a specific focus on their applications in process monitoring and other related areas.
Liu and Xie (2020) [813]	Provides a review of the application of kernel techniques for constructing data-driven soft sensors.
Curreri et al. (2020) [814]	Reviews variable selection methods utilized in the development of soft sensors.
Jiao et al. (2020) [815]	Provides a review of the application of machine learning methods for ensuring the health and safety of processes.
Arunthavanathan et al. (2021) [816]	Presents an overview of three essential aspects for ensuring process safety: fault detection and diagnosis, risk assessment, and management of abnormal situations.
Taqvi et al. (2021) [817]	Reviews various methods with a specific emphasis on the distinction between supervised and unsupervised approaches.

Table A1. *Cont.*

Reference	Description
Okada et al. (2021) [655]	Divides the approaches into signal analysis, model-based, data-based, and hybrid methodologies.
Sun and Ge (2021) [818]	Provides a review of deep learning methods for applications in soft sensors.
Pani (2022) [102]	Reviews extensions of the PCA model using the kernel technique for nonlinear monitoring.
Zhao (2022) [3]	Reviews methods for non-stationary monitoring and provides perspectives for future developments in the field.
Wang et al. (2022) [819]	Reviews recursive techniques with a specific focus on multivariate statistical methods.
Qian et al. (2022) [340]	Provides a review of the application of autoencoders for representation learning in the context of process monitoring problems.
Bi et al. (2022) [654]	Presents a comprehensive review of fault detection and diagnosis techniques, along with perspectives and challenges for future development in the field.
Nawaz et al. (2022) [124]	Applies bibliometric techniques to conduct a focused review of strategies adopted for multiscale monitoring.
Kong et al. (2022) [2]	Reviews the application of multivariate statistical methods in the context of big data.
Webert et al. (2022) [820]	Reviews strategies for fault handling in the context of detection, classification, and prioritization activities.
Melo et al. (2022) [5]	Reviews open benchmarks for the assessment of process monitoring techniques, conducting an extensive exploratory analysis and critical discussion of their effectiveness and limitations.
Perera et al. (2023) [663]	Presents a critical review of the industrial application of soft sensors based on artificial intelligence.
Yan et al. (2023) [662]	Reviews methods for real-time fault diagnosis applied to equipment and process monitoring, with a specific focus on machine learning techniques.
Palla and Pani (2023) [263]	Reviews the application of the ICA model to process monitoring problems.
Yu and Zhang (2023) [1]	Reviews the application of deep learning models to process monitoring problems.
Liu et al. (2023) [821]	Presents a review and perspectives on the management of abnormal situations in chemical processes.
Ramírez-Sanz et al. (2023) [822]	Reviews works focused on semi-supervised methodologies and presents best practices for applications in this context.
Lou et al. (2023) [823]	Reviews the application of deep learning models for fault diagnosis and health monitoring.

Table A2. Comparison articles about process monitoring in PSE literature, in chronological order.

Reference	Description
Westerhuis et al. (1999) [65]	Compares various approaches to batch process monitoring.
Tien et al. (2004) [824]	Compares variations in the PCA model, including conventional PCA, MPCA, APCA and EWPCA.
Alcala and Qin (2011) [56]	Analyzes five commonly used diagnostic methods and proposes their unification into three general methods, providing a comprehensive framework for diagnosis in various applications.
Yin et al. (2012) [825]	Compares the applications of various multivariate statistical methods to the TEP benchmark.
Zhang et al. (2015) [826]	Compares the performances of KPI-based monitoring methods.
Jing and Hou (2015) [827]	Compares the applications of PCA and SVM methods.
Li and Qin (2016) [48]	Compares different monitoring schemes for processes with non-Gaussian distributions.
Askarian et al. (2016) [828]	Compares different approaches for handling missing data, providing an analysis of their performances in fault diagnosis applications.

Table A2. *Cont.*

Reference	Description
Rato et al. (2016) [829]	Compares different PCA approaches for handling large-scale, time-dependent processes.
Zhang et al. (2017) [830]	Compares the performance of statistics T^2 and Q for detection of additive and multiplicative faults.
Fuentes-García et al. (2018) [50]	Proposes a methodology for conducting experimental comparisons between multiple diagnostic frameworks in the context of applying the PCA model.
Zhang et al. (2018) [160]	Compares the applications of PCA, CCA, and PLS, along with their dynamic extensions.
Vidal-Puig et al. (2019) [58]	Compares diagnostic schemes applied to multivariate statistical models, specifically PCA and PLS.
Shrivastava (2021) [492]	Compares different strategies for ensembling models based on decision trees.
Lomov et al. (2021) [410]	Compares the applications of different deep learning architectures.
Fernandes et al. (2022) [831]	Compares dynamic models for multivariate statistical process monitoring.
Zheng et al. (2022) [832]	Compares linear dynamic models for multivariate statisticaç process monitoring.
Hansen et al. (2023) [833]	Compares PCA and deep learning models applied for fault detection in power plants.

Table A3. Books about process monitoring in PSE literature, in chronological order.

Reference	Description
Fault Detection and Diagnosis in Chemical and Petrochemical Processes [834]	Describes applications in both data-driven and model-based approaches, including the use of control charts, pattern recognition techniques, and parameter estimation methods.
Data Mining and Knowledge Discovery for Process Monitoring and Control [835]	Covers a range of topics including signal processing techniques, feature extraction methods, identification of operational states, and other related areas.
Fault Detection and Diagnosis in Industrial Systems [20]	Describes multivariate statistical methods such as PCA, FDA, PLS, and CVA, and applies them to the Tennessee Eastman model. In the final chapters, the author focuses on analytical methods and knowledge-based methods.
Multivariate Statistical Process Control with Industrial Applications [836]	Describes various approaches for applying T^2 statistics in the field of process monitoring.
Statistical Monitoring of Complex Multivariate Processes: With Applications in Industrial Process Control [27]	Provides a detailed description of PCA and PLS methods, covering their fundamental principles and exploring recent advancements in these techniques.
Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods [314]	Presents the monitoring problem from the perspective of the machine learning field, exploring various approaches such as neural networks, statistical learning theory, kernel methods, tree methods, and other relevant techniques.
Multivariate Statistical Process Control: Process Monitoring Methods and Applications [837]	Similar to the review published in the same year by two of the same authors [800], the book classifies and presents methods based on the type of process, categorizing them into continuous, batch, non-Gaussian, nonlinear, time-variant/multimodal, and dynamic processes.
Data-driven Design of Fault Diagnosis and Fault-tolerant Control Systems [164]	Addresses three main themes: (i) basic methods of multivariate analysis applied to the monitoring problem; (ii) the data-driven design of input/output data models and observer-based fault detection and isolation schemes; and (iii) the data-driven design of fault-tolerant control systems, with a particular emphasis on managing the lifetime of control systems.

Table A3. *Cont.*

Reference	Description
Chemometric Monitoring: Product Quality Assessment, Process Fault Detection, and Applications [47]	Focused on the context of chemometrics, the book explores a range of techniques including PCA, PLS, SVM, clustering, and more. In addition to covering these methods, it also addresses topics that may not have been extensively covered in previous books, such as data generation and pre-processing. The book provides practical illustrations with numerous real-world application examples.
Advanced methods for fault diagnosis and fault-tolerant control [838]	Covers both data-based and model-based techniques, with a specific emphasis on the application of machine learning methods in the context of fault-tolerant control.

Appendix B. List of Commercial Tools Available for Industrial Process Monitoring in the Context of PSE

- PLS_Toolbox: Developed in 1991 by Barry Wise, who introduced the application of the PCA technique for process monitoring [11], the PLS_Toolbox stands as a pioneering and widely utilized software in the field of chemometrics [839–842]. The toolbox encompasses a broad range of techniques tailored to address chemometrics problems, including exploratory analysis, classification, regression, and process monitoring. It works as the toolbox of the Matlab computing environment and is sold by the company Eigenvector Research. URL: http://eigenvector.com/software/pls-toolbox/ (accessed on 16 January 2024).
- AspenTech Suite: the company AspenTech offers a range of solutions, including:
  - AspenOne Process Explorer: Provides real-time visualization, analysis, and monitoring of plant operation data. URL: https://www.aspentech.com/en/products/dataworks/aspenone-process-explorer (accessed on 16 January 2024).
  - Aspen ProMV: Enables the creation of alerts regarding process health, helping to identify potential issues and ensure smooth operations. URL: https://www.aspentech.com/en/products/apm/aspen-promv (accessed on 16 January 2024).
  - Aspen OnLine: Offers the possibility to calibrate predictive models to monitor processes in real time together with process simulations. URL: https://www.aspentech.com/en/products/engineering/aspen-online (accessed on 16 January 2024).
  - Aspen Process Pulse: Enables the generation of warnings about problems related to product quality with optimization goals. URL: https://www.aspentech.com/en/products/apm/aspen-process-pulse (accessed on 16 January 2024).
  - Aspen Mtell: Focuses on predictive maintenance of equipment. URL: https://www.aspentech.com/en/products/apm/aspen-mtell (accessed on 16 January 2024).
- Aveva Insight: Marketed by the company Aveva, this software enables remote, real-time monitoring of production assets. URL: https://www.aveva.com/en/products/insight/ (accessed on 16 January 2024).
- Precognize: The company Precognize provides a homonymous software with specific solutions for predictive maintenance and process monitoring. URL: https://www.precog.co/ (accessed on 16 January 2024).
- Honeywell Forge: Offered by the company Honeywell, this platform brings together several solutions in the field of digital transformation, which include monitoring, diagnosing and predicting the performance of process plants. URL: https://www.honeywellforge.ai (accessed on 16 January 2024).
- IBM Maximo Suite: Marketed by the company IBM, this platform offers solutions for predictive maintenance based on process condition analysis. URL: https://www.ibm.com/products/maximo/predictive-maintenance (accessed on 16 January 2024).

- Braincube Suite: The company Braincube offers several applications for data visualization, process condition monitoring and quality control. URL: https://braincube.com/solutions/ (accessed on 16 January 2024).
- Seeq Suite: The company Seeq provides a platform for time series analysis that includes several stages of the process monitoring pipeline, such as exploratory data analysis, predictive modeling and alarm generation. URL: https://www.seeq.com/product (accessed on 16 January 2024).
- Proficy CSense: Created by the company GE, this tool enables the identification of issues, root cause analysis, prediction of future performance, and automation of actions. URL: https://www.ge.com/digital/applications/proficy-csense (accessed on 16 January 2024).
- JMP Software: The company JMP offers a set of solutions for the entire industrial data analysis flow, including statistical monitoring. URL: https://www.jmp.com/en_us/software.html (accessed on 16 January 2024).
- Visplore: Marketed by the company Visplore GmbH, this software provides solutions for exploratory data analysis, including visualization and correlation analysis, root cause search, regression and comparison. URL: https://visplore.com/ (accessed on 16 January 2024).
- Statistica: Developed by the company TIBCO, this tool has a module for multivariate statistical process control that makes available the PCA and PLS techniques, as well as some of its variants. URL: https://community.tibco.com/s/article/TIBCO-Statistica-Multivariate-Statistical-Process-Control (accessed on 16 January 2024).

References

1. Yu, J.; Zhang, Y. Challenges and opportunities of deep learning-based process fault detection and diagnosis: A review. *Neural Comput. Appl.* **2023**, *35*, 211–252. [CrossRef]
2. Kong, X.; Jiang, X.; Zhang, B.; Yuan, J.; Ge, Z. Latent variable models in the era of industrial big data: Extension and beyond. *Annu. Rev. Control* **2022**, *54*, 167–199. [CrossRef]
3. Zhao, C. Perspectives on nonstationary process monitoring in the era of industrial artificial intelligence. *J. Process Control* **2022**, *116*, 255–272. [CrossRef]
4. Peres, F.A.P.; Fogliatto, F.S. Variable selection methods in multivariate statistical process control: A systematic literature review. *Comput. Ind. Eng.* **2018**, *115*, 603–619. [CrossRef]
5. Melo, A.; Câmara, M.M.; Clavijo, N.; Pinto, J.C. Open benchmarks for assessment of process monitoring and fault diagnosis techniques: A review and critical analysis. *Comput. Chem. Eng.* **2022**, *165*, 107964. [CrossRef]
6. Shewhart, W.A. Quality control charts. *Bell Syst. Tech. J.* **1926**, *5*, 593–603. [CrossRef]
7. Shewhart, W.A. *Economic Control of Quality of Manufactured Product*; D. Van Nostrand Company, Inc.: Lancaster, UK, 1931.
8. Montgomery, D.C. *Introduction to Statistical Quality Control*, 7th ed.; Wiley: Hoboken, NJ, USA, 2012.
9. Hotelling, H. Multivariate quality control illustrated by the air testing of sample bombsites. In *Techniques of Statistical Analysis*; Eisenhart, C., Hastay, M., Wallis, W., Eds.; McGraw-Hill: New York, NY, USA, 1947.
10. Das, A.; Maiti, J.; Banerjee, R. Process monitoring and fault detection strategies: A review. *Int. J. Qual. Reliab. Manag.* **2012**, *29*, 720–752. [CrossRef]
11. Wise, B.; Veltkamp, D.; Davis, B.; Ricker, N.; Kowalski, B. Principal components analysis for monitoring the West Valley Liquid Fed Ceramic Melter. *Waste Manag.* **1988**, *88*, 811–818.
12. MacGregor, J.F.; Kourti, T. Statistical process control of multivariate processes. *Control Eng. Pract.* **1995**, *3*, 403–414. [CrossRef]
13. Wise, B.M.; Gallagher, N.B. The process chemometrics approach to process monitoring and fault detection. *J. Process Control* **1996**, *6*, 329–348. [CrossRef]
14. Qin, S.J. Statistical process monitoring: Basics and beyond. *J. Chemom.* **2003**, *17*, 480–502. [CrossRef]
15. Md Nor, N.; Che Hassan, C.R.; Hussain, M.A. A review of data-driven fault detection and diagnosis methods: applications in chemical process systems. *Rev. Chem. Eng.* **2020**, *36*, 513–553. [CrossRef]
16. Qin, S.J.; Chiang, L.H. Advances and opportunities in machine learning for process data analytics. *Comput. Chem. Eng.* **2019**, *126*, 465–473. [CrossRef]
17. Weese, M.; Martinez, W.; Megahed, F.M.; Jones-Farmer, L.A. Statistical Learning Methods Applied to Process Monitoring: An Overview and Perspective. *J. Qual. Technol.* **2016**, *48*, 4–24. [CrossRef]
18. Qin, S.J. Survey on data-driven industrial process monitoring and diagnosis. *Annu. Rev. Control* **2012**, *36*, 220–234. [CrossRef]
19. Qin, S.J. Process data analytics in the era of big data. *AIChE J.* **2014**, *60*, 3092–3100. [CrossRef]
20. Chiang, L.H.; Russell, E.L.; Braatz, R.D. *Fault Detection and Diagnosis in Industrial Systems*; Springer: London, UK, 2001.

21. Rencher, A.C.; Christensen, W.F. *Methods of Multivariate Analysis*, 3rd ed.; Wiley Series in Probability and Statistics; Wiley: Hoboken, NJ, USA, 2012.
22. Borga, M.; Landelius, T.; Knutsson, H. *A Unified Approach to PCA, PLS, MLR and CCA*; Linköping University: Linköping, Sweden, 1997.
23. Blei, D.M. Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models. *Annu. Rev. Stat. Its Appl.* **2014**, *1*, 203–232. [CrossRef]
24. Kourti, T. Process analysis and abnormal situation detection: From theory to practice. *IEEE Control Syst.* **2002**, *22*, 10–25. [CrossRef]
25. Qin, S.J. Data-driven Fault Detection and Diagnosis for Complex Industrial Processes. *IFAC Proc. Vol.* **2009**, *42*, 1115–1125. [CrossRef]
26. MacGregor, J.; Cinar, A. Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven methods. *Comput. Chem. Eng.* **2012**, *47*, 111–120. [CrossRef]
27. Kruger, U.; Xie, L. *Statistical Monitoring of Complex Multivariate Processes: With Applications in Industrial Process Control*; John Wiley & Sons, Ltd.: Chichester, UK, 2012. [CrossRef]
28. Yin, S.; Ding, S.X.; Xie, X.; Luo, H. A Review on Basic Data-Driven Approaches for Industrial Process Monitoring. *IEEE Trans. Ind. Electron.* **2014**, *61*, 6418–6428. [CrossRef]
29. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer Series in Statistics; Springer: New York, NY, USA, 2009.
30. Prata, D.M.; Lima, E.L.; Pinto, J.C. Nonlinear Dynamic Data Reconciliation in Real Time in Actual Processes. In *Computer Aided Chemical Engineering*; Elsevier: Amsterdam, The Netherlands, 2009; Volume 27, pp. 47–54. [CrossRef]
31. Câmara, M.M.; Soares, R.M.; Feital, T.; Anzai, T.K.; Diehl, F.C.; Thompson, P.H.; Pinto, J.C. Numerical Aspects of Data Reconciliation in Industrial Applications. *Processes* **2017**, *5*, 56. [CrossRef]
32. Menezes, D.Q.F.d.; de Sá, M.C.C.; Fontoura, T.B.; Anzai, T.K.; Diehl, F.C.; Thompson, P.H.; Pinto, J.C. Modeling of Spiral Wound Membranes for Gas Separations—Part II: Data Reconciliation for Online Monitoring. *Processes* **2020**, *8*, 1035. [CrossRef]
33. Andrade, G.M.; De Menezes, D.Q.; Soares, R.M.; Lemos, T.S.; Teixeira, A.F.; Ribeiro, L.D.; Vieira, B.F.; Pinto, J.C. Virtual flow metering of production flow rates of individual wells in oil and gas platforms through data reconciliation. *J. Pet. Sci. Eng.* **2022**, *208*, 109772. [CrossRef]
34. Pearson, K. LIII—On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [CrossRef]
35. Izenman, A.J. *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*; Springer Texts in Statistics; Springer: New York, NY, USA; London, UK, 2008.
36. Shawe-Taylor, J.; Cristianini, N. Pattern analysis using eigen-decompositions. In *Kernel Methods for Pattern Analysis*; Cambridge University Press: Cambridge, UK, 2004; pp. 140–194.
37. De Bie, T.; Cristianini, N.; Rosipal, R. Eigenproblems in Pattern Recognition. In *Handbook of Geometric Computing*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 129–167. [CrossRef]
38. Shlens, J. A Tutorial on Principal Component Analysis, v3.0.2. *arXiv* **2014**, arXiv:1404.1100. [CrossRef]
39. Himes, D.M.; Storer, R.H.; Georgakis, C. Determination of the number of principal components for disturbance detection and isolation. In Proceedings of the 1994 American Control Conference—ACC'94, Baltimore, MD, USA, 29 June–1 July 1994; Volume 2, pp. 1279–1283. [CrossRef]
40. Valle, S.; Li, W.; Qin, S.J. Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods. *Ind. Eng. Chem. Res.* **1999**, *38*, 4389–4401. [CrossRef]
41. Qin, S.J.; Dunia, R. Determining the number of principal components for best reconstruction. *J. Process Control* **2000**, *10*, 245–250. [CrossRef]
42. Tamura, M.; Tsujita, S. A study on the number of principal components and sensitivity of fault detection using PCA. *Comput. Chem. Eng.* **2007**, *31*, 1035–1046. [CrossRef]
43. Saccenti, E.; Camacho, J. Determining the number of components in principal components analysis: A comparison of statistical, crossvalidation and approximated methods. *Chemom. Intell. Lab. Syst.* **2015**, *149*, 99–116. [CrossRef]
44. Wu, D.; Zhou, D.; Chen, M. Performance-Driven Component Selection in the Framework of PCA for Process Monitoring: A Dynamic Selection Approach. *IEEE Trans. Control Syst. Technol.* **2022**, *30*, 1171–1185. [CrossRef]
45. Ku, W.; Storer, R.H.; Georgakis, C. Disturbance detection and isolation by dynamic principal component analysis. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 179–196. [CrossRef]
46. Ketelaere, B.; Hubert, M.; Schmitt, E. Overview of PCA-Based Statistical Process-Monitoring Methods for Time-Dependent, High-Dimensional Data. *J. Qual. Technol.* **2015**, *47*, 318–335. [CrossRef]
47. Kundu, M.; Kundu, P.K.; Damarla, S.K. *Chemometric Monitoring: Product Quality Assessment, Process Fault Detection, and Applications*; CRC Press: Boca Raton, FL, USA, 2017. [CrossRef]
48. Li, G.; Qin, S.J. Comparative study on monitoring schemes for non-Gaussian distributed processes. *J. Process Control* **2016**, *67*, 69–82. [CrossRef]
49. Downs, J.; Vogel, E. A plant-wide industrial process control problem. *Comput. Chem. Eng.* **1993**, *17*, 245–255. [CrossRef]

50. Fuentes-García, M.; Maciá-Fernández, G.; Camacho, J. Evaluation of diagnosis methods in PCA-based Multivariate Statistical Process Control. *Chemom. Intell. Lab. Syst.* **2018**, *172*, 194–210. [CrossRef]
51. Milman, O. Why does PCA assume Gaussian Distribution? 2018. Available online: https://datascience.stackexchange.com/a/37178 (accessed on 16 January 2024).
52. Shlens, J. A Tutorial on Principal Component Analysis, v2.0. 2005. Available online: https://www.cs.cmu.edu/~elaw/papers/pca.pdf (accessed on 16 January 2024).
53. Rodgers, J.L.; Nicewander, W.A.; Toothaker, L. Linearly Independent, Orthogonal, and Uncorrelated Variables. *Am. Stat.* **1984**, *38*, 133–134.
54. Kourti, T. Pharmaceutical Manufacturing: The Role of Multivariate Analysis in Design Space, Control Strategy, Process Understanding, Troubleshooting, and Optmization. In *Chemical Engineering in the Pharmaceutical Industry*; am Ende, D.J., am Ende, M.T., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2019; pp. 601–629. [CrossRef]
55. Jackson, J.E.; Mudholkar, G.S. Control Procedures for Residuals Associated with Principal Component Analysis. *Technometrics* **1979**, *21*, 341–349. [CrossRef]
56. Alcala, C.F.; Qin, S.J. Analysis and generalization of fault diagnosis methods for process monitoring. *J. Process Control* **2011**, *21*, 322–330. [CrossRef]
57. Alcala, C.F.; Qin, S.J. Reconstruction-based contribution for process monitoring. *Automatica* **2009**, *45*, 1593–1600. [CrossRef]
58. Vidal-Puig, S.; Vitale, R.; Ferrer, A. Data-driven supervised fault diagnosis methods based on latent variable models: A comparative study. *Chemom. Intell. Lab. Syst.* **2019**, *187*, 41–52. [CrossRef]
59. Westerhuis, J.A.; Gurden, S.P.; Smilde, A.K. Generalized contribution plots in multivariate statistical process monitoring. *Chemom. Intell. Lab. Syst.* **2000**, *51*, 95–114. [CrossRef]
60. Van den Kerkhof, P.; Vanlaer, J.; Gins, G.; Van Impe, J.F. Analysis of smearing-out in contribution plot based fault isolation for Statistical Process Control. *Chem. Eng. Sci.* **2013**, *104*, 285–293. [CrossRef]
61. Alakent, B. Reducing smearing effect in contribution plots and improving fault detection via polynomial approximated isomap embeddings. *Can. J. Chem. Eng.* **2023**, *101*, 2768–2789. [CrossRef]
62. Nomikos, P.; MacGregor, J.F. Monitoring batch processes using multiway principal component analysis. *AIChE J.* **1994**, *40*, 1361–1375. [CrossRef]
63. Kosanovich, K.; Piovoso, M.; Dahl, K.; MacGregor, J.; Nomikos, P. Multi-way PCA applied to an industrial batch process. In Proceedings of the 1994 American Control Conference—ACC '94, Baltimore, MD, USA, 29 June–1 July 1994; Volume 2, pp. 1294–1298. [CrossRef]
64. Wise, B.M.; Gallagher, N.B.; Butler, S.W.; White, D.D.; Barna, G.G. A comparison of principal component analysis, multiway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process. *J. Chemom.* **1999**, *13*, 379–396. [CrossRef]
65. Westerhuis, J.A.; Kourti, T.; MacGregor, J.F. Comparing alternative approaches for multivariate statistical analysis of batch process data. *J. Chemom.* **1999**, *13*, 397–413. [CrossRef]
66. Camacho, J.; Picó, J.; Ferrer, A. The best approaches in the on-line monitoring of batch processes based on PCA: Does the modelling structure matter? *Anal. Chim. Acta* **2009**, *642*, 59–68. [CrossRef]
67. Rendall, R.; Chiang, L.H.; Reis, M.S. Data-driven methods for batch data analysis—A critical overview and mapping on the complexity scale. *Comput. Chem. Eng.* **2019**, *124*, 1–13. [CrossRef]
68. Yao, Y.; Gao, F. A survey on multistage/multiphase statistical modeling methods for batch processes. *Annu. Rev. Control* **2009**, *33*, 172–183. [CrossRef]
69. Zhu, J.; Yao, Y.; Gao, F. Multiphase two-dimensional time-slice dynamic system for batch process monitoring. *J. Process Control* **2020**, *85*, 184–198. [CrossRef]
70. Lu, N.; Yao, Y.; Gao, F.; Wang, F. Two-dimensional dynamic PCA for batch process monitoring. *AIChE J.* **2005**, *51*, 3300–3304. [CrossRef]
71. Dong, Y.; Qin, S.J. A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. *J. Process Control* **2018**, *67*, 1–11. [CrossRef]
72. Li, G.; Liu, B.; Qin, S.J.; Zhou, D. Dynamic latent variable modeling for statistical process monitoring. *IFAC Proc. Vol.* **2011**, *44*, 12886–12891. [CrossRef]
73. Li, G.; Qin, S.J.; Zhou, D. A New Method of Dynamic Latent-Variable Modeling for Process Monitoring. *IEEE Trans. Ind. Electron.* **2014**, *61*, 6438–6445. [CrossRef]
74. Dong, Y.; Qin, S.J. Dynamic latent variable analytics for process operations and control. *Comput. Chem. Eng.* **2018**, *114*, 69–80. [CrossRef]
75. Guo, L.; Wu, P.; Lou, S.; Gao, J.; Liu, Y. A multi-feature extraction technique based on principal component analysis for nonlinear dynamic process monitoring. *J. Process Control* **2020**, *85*, 159–172. [CrossRef]
76. Lu, S.; Huang, S. Segmentation of Multivariate Industrial Time Series Data Based on Dynamic Latent Variable Predictability. *IEEE Access* **2020**, *8*, 112092–112103. [CrossRef]
77. Wold, S. Exponentially weighted moving principal components analysis and projections to latent structures. *Chemom. Intell. Lab. Syst.* **1994**, *23*, 149–161. [CrossRef]

78. Gallagher, N.B.; Wise, B.M.; Butler, S.W.; White, D.D.; Barna, G.G. Development and Benchmarking of Multivariate Statistical Process Control Tools for a Semiconductor Etch Process: Improving Robustness through Model Updating. *IFAC Proc. Vol.* **1997**, *30*, 79–84. [CrossRef]
79. Li, W.; Yue, H.H.; Valle-Cervantes, S.; Qin, S.J. Recursive PCA for adaptive process monitoring. *J. Process Control* **2000**, *10*, 471–486. [CrossRef]
80. Jeng, J.C. Adaptive process monitoring using efficient recursive PCA and moving window PCA algorithms. *J. Taiwan Inst. Chem. Eng.* **2010**, *41*, 475–481. [CrossRef]
81. Aldrich, C. Process Fault Diagnosis for Continuous Dynamic Systems Over Multivariate Time Series. In *Time Series Analysis—Data, Methods, and Applications*; IntechOpen: Rijeka, Croatia, 2019. [CrossRef]
82. Kramer, M.A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **1991**, *37*, 233–243. [CrossRef]
83. Dong, D.; McAvoy, T. Nonlinear principal component analysis—Based on principal curves and neural networks. *Comput. Chem. Eng.* **1996**, *20*, 65–78. [CrossRef]
84. Hastie, T.; Stuetzle, W. Principal Curves. *J. Am. Stat. Assoc.* **1989**, *84*, 502–516. [CrossRef]
85. Jia, F.; Martin, E.B.; Morris, A.J. Non-linear principal components analysis with application to process fault detection. *Int. J. Syst. Sci.* **2000**, *31*, 1473–1487. [CrossRef]
86. Tan, S.; Mayrovouniotis, M.L. Reducing data dimensionality through optimizing neural network inputs. *AIChE J.* **1995**, *41*, 1471–1480. [CrossRef]
87. Liu, X.; Li, K.; McAfee, M.; Irwin, G.W. Improved nonlinear PCA for process monitoring using support vector data description. *J. Process Control* **2011**, *21*, 1306–1317. [CrossRef]
88. Scholkopf, B.; Smola, A.; Müller, K.R. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Comput.* **1998**, *10*, 1299–1319. [CrossRef]
89. Lee, J.M.; Yoo, C.; Choi, S.W.; Vanrolleghem, P.A.; Lee, I.B. Nonlinear process monitoring using kernel principal component analysis. *Chem. Eng. Sci.* **2004**, *59*, 223–234. [CrossRef]
90. Choi, S.W.; Lee, C.; Lee, J.M.; Park, J.H.; Lee, I.B. Fault detection and identification of nonlinear processes based on kernel PCA. *Chemom. Intell. Lab. Syst.* **2005**, *75*, 55–67. [CrossRef]
91. Deng, X.; Tian, X.; Chen, S. Modified kernel principal component analysis based on local structure analysis and its application to nonlinear process fault diagnosis. *Chemom. Intell. Lab. Syst.* **2013**, *127*, 195–209. [CrossRef]
92. Yao, M.; Wang, H. On-line monitoring of batch processes using generalized additive kernel principal component analysis. *J. Process Control* **2015**, *28*, 56–72. [CrossRef]
93. Deng, X.; Tian, X.; Chen, S.; Harris, C.J. Deep Principal Component Analysis Based on Layerwise Feature Extraction and Its Application to Nonlinear Process Monitoring. *IEEE Trans. Control Syst. Technol.* **2018**, *27*, 2526–2540. [CrossRef]
94. Deng, X.; Tian, X.; Chen, S.; Harris, C.J. Nonlinear Process Fault Diagnosis Based on Serial Principal Component Analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 560–572. [CrossRef]
95. Jiang, Q.; Yan, X. Parallel PCA–KPCA for nonlinear process monitoring. *Control Eng. Pract.* **2018**, *80*, 17–25. [CrossRef]
96. Jiang, Q.; Yan, X. Improved fault detection in nonlinear chemical processes using WKPCA-SVDD. *Korean J. Chem. Eng.* **2014**, *31*, 1935–1942. [CrossRef]
97. Wang, K.; Yan, C.; Yuan, X.; Wang, Y.; Liu, C. A reduced nonstationary discrete convolution kernel for multimode process monitoring. *Int. J. Mach. Learn. Cybern.* **2022**, *13*, 3711–3725. [CrossRef]
98. Bernal-de Lázaro, J.M.; Cruz-Corona, C.; Silva-Neto, A.J.; Llanes-Santiago, O. Criteria for optimizing kernel methods in fault monitoring process: A survey. *ISA Trans.* **2022**, *127*, 259–272. [CrossRef]
99. Apsemidis, A.; Psarakis, S.; Moguerza, J.M. A review of machine learning kernel methods in statistical process monitoring. *Comput. Ind. Eng.* **2020**, *142*, 106376. [CrossRef]
100. Hiden, H. Non-linear principal components analysis using genetic programming. In Proceedings of the Second International Conference on Genetic Algorithms in Engineering Systems, Glasgow, UK, 2–4 September 1997; Volume 1997, pp. 302–307. [CrossRef]
101. Pilario, K.E.; Shafiee, M.; Cao, Y.; Lao, L.; Yang, S.H. A Review of Kernel Methods for Feature Extraction in Nonlinear Process Monitoring. *Processes* **2019**, *8*, 24. [CrossRef]
102. Pani, A.K. Non-linear process monitoring using kernel principal component analysis: A review of the basic and modified techniques with industrial applications. *Braz. J. Chem. Eng.* **2022**, *39*, 327–344. [CrossRef]
103. Quiñones-Grueiro, M.; Prieto-Moreno, A.; Verde, C.; Llanes-Santiago, O. Data-driven monitoring of multimode continuous processes: A review. *Chemom. Intell. Lab. Syst.* **2019**, *189*, 56–71. [CrossRef]
104. Feital, T.; Kruger, U.; Dutra, J.; Pinto, J.C.; Lima, E.L. Modeling and performance monitoring of multivariate multimodal processes. *AIChE J.* **2013**, *59*, 1557–1569. [CrossRef]
105. Choi, S.W.; Park, J.H.; Lee, I.B. Process monitoring using a Gaussian mixture model via principal component analysis and discriminant analysis. *Comput. Chem. Eng.* **2004**, *28*, 1377–1387. [CrossRef]
106. Chen, T.; Morris, J.; Martin, E. Probability density estimation via an infinite Gaussian mixture model: Application to statistical process monitoring. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2006**, *55*, 699–715. [CrossRef]

107. Chen, T.; Zhang, J. On-line multivariate statistical monitoring of batch processes using Gaussian mixture model. *Comput. Chem. Eng.* **2010**, *34*, 500–507. [CrossRef]
108. Lyu, Y.; Zhou, L.; Cong, Y.; Zheng, H.; Song, Z. Multirate Mixture Probability Principal Component Analysis for Process Monitoring in Multimode Processes. *IEEE Trans. Autom. Sci. Eng.* **2023**, 1–12. [CrossRef]
109. Bakshi, B.R.; Locher, G.; Stephanopoulos, G.; Stephanopoulous, G. Analysis of operating data for evaluation, diagnosis and control of batch operations. *J. Process Control* **1994**, *4*, 179–194. [CrossRef]
110. Rioul, O.; Vetterli, M. Wavelets and signal processing. *Signal Process. Mag. IEEE* **1991**, *8*, 14–38. [CrossRef]
111. Kosanovich, K.A. Improvements in the Development of Models for Monitoring & Detection. *IFAC Proc. Vol.* **1997**, *30*, 29–34. [CrossRef]
112. Kosanovich, K.A.; Piovoso, M.J. PCA of wavelet transformed process data for monitoring. *Intell. Data Anal.* **1997**, *1*, 85–99. [CrossRef]
113. Bakshi, B.R. Multiscale PCA with application to multivariate statistical process monitoring. *AIChE J.* **1998**, *44*, 1596–1610. [CrossRef]
114. Yoon, S.; MacGregor, J.F. Unifying PCA and multiscale approaches to fault detection and isolation. *IFAC Proc. Vol.* **2001**, *34*, 433–438. [CrossRef]
115. Misra, M.; Yue, H.; Qin, S.; Ling, C. Multivariate process monitoring and fault diagnosis by multi-scale PCA. *Comput. Chem. Eng.* **2002**, *26*, 1281–1293. [CrossRef]
116. Aradhye, H.B.; Bakshi, B.R.; Strauss, R.A.; Davis, J.F. Multiscale SPC using wavelets: Theoretical analysis and properties. *AIChE J.* **2003**, *49*, 939–958. [CrossRef]
117. Yoon, S.; MacGregor, J.F. Principal-component analysis of multiscale data for process monitoring and fault diagnosis. *AIChE J.* **2004**, *50*, 2891–2903. [CrossRef]
118. Reis, M.S.; Saraiva, P.M. Multiscale statistical process control with multiresolution data. *AIChE J.* **2006**, *52*, 2107–2119. [CrossRef]
119. Zhang, Y.; Ma, C. Fault diagnosis of nonlinear processes using multiscale KPCA and multiscale KPLS. *Chem. Eng. Sci.* **2011**, *66*, 64–72. [CrossRef]
120. Alawi, A.; Zhang, J.; Morris, J. Multiscale Multiblock Batch Monitoring: Sensor and Process Drift and Degradation. *Org. Process Res. Dev.* **2015**, *19*, 145–157. [CrossRef]
121. Sheriff, M.Z.; Mansouri, M.; Karim, M.N.; Nounou, H.; Nounou, M. Fault detection using multiscale PCA-based moving window GLRT. *J. Process Control* **2017**, *54*, 47–64. [CrossRef]
122. Nawaz, M.; Maulud, A.S.; Zabiri, H.; Suleman, H.; Tufa, L.D. Multiscale Framework for Real-Time Process Monitoring of Nonlinear Chemical Process Systems. *Ind. Eng. Chem. Res.* **2020**, *59*, 18595–18606. [CrossRef]
123. Reis, M. Multiscale and Multi-Granularity Process Analytics: A Review. *Processes* **2019**, *7*, 61. [CrossRef]
124. Nawaz, M.; Maulud, A.S.; Zabiri, H. Analysis of multiscale process monitoring in industrial processes from a bibliometric perspective. *Comput. Chem. Eng.* **2022**, *168*, 108055. [CrossRef]
125. Qin, S.J.; Valle, S.; Piovoso, M.J. On unifying multiblock analysis with application to decentralized process monitoring. *J. Chemom.* **2001**, *15*, 715–742. [CrossRef]
126. Huang, J.; Yan, X. Relevant and independent multi-block approach for plant-wide process and quality-related monitoring based on KPCA and SVDD. *ISA Trans.* **2018**, *73*, 257–267. [CrossRef] [PubMed]
127. Ge, Z.; Song, Z. Distributed PCA Model for Plant-Wide Process Monitoring. *Ind. Eng. Chem. Res.* **2013**, *52*, 1947–1957. [CrossRef]
128. Jiang, Q.; Yan, X. Nonlinear plant-wide process monitoring using MI-spectral clustering and Bayesian inference-based multiblock KPCA. *J. Process Control* **2015**, *32*, 38–50. [CrossRef]
129. Cao, Y.; Yuan, X.; Wang, Y.; Gui, W. Hierarchical hybrid distributed PCA for plant-wide monitoring of chemical processes. *Control Eng. Pract.* **2021**, *111*, 104784. [CrossRef]
130. Ge, Z. Review on data-driven modeling and monitoring for plant-wide industrial processes. *Chemom. Intell. Lab. Syst.* **2017**, *171*, 16–25. [CrossRef]
131. Jiang, Q.; Yan, X.; Huang, B. Review and Perspectives of Data-Driven Distributed Monitoring for Industrial Plant-Wide Processes. *Ind. Eng. Chem. Res.* **2019**, *58*, 12899–12912. [CrossRef]
132. Feital, T.d.S. Monitoramento da Condição de Processos Químicos Industriais. Ph.D. Thesis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, 2011.
133. Monteiro, A.D.R.D.; Feital, T.D.S.; Pinto, J.C. Statistical Aspects of Near-Infrared Spectroscopy for the Characterization of Errors and Model Building. *Appl. Spectrosc.* **2017**, *71*, 1665–1676. [CrossRef] [PubMed]
134. Tipping, M.E.; Bishop, C.M. Probabilistic principal component analysis. *J. R. Stat. Soc.* **1999**, *61*, 611–622. [CrossRef]
135. Ge, Z.; Song, Z. Mixture Bayesian regularization method of PPCA for multimode process monitoring. *AIChE J.* **2010**, *56*, 2838–2849. [CrossRef]
136. Feital, T.; Kruger, U.; Xie, L.; Schubert, U.; Lima, E.L.; Pinto, J.C. A unified statistical framework for monitoring multivariate systems with unknown source and error signals. *Chemom. Intell. Lab. Syst.* **2010**, *104*, 223–232. [CrossRef]
137. Narasimhan, S.; Shah, S.L. Model identification and error covariance matrix estimation from noisy data using PCA. *Control Eng. Pract.* **2008**, *16*, 146–155. [CrossRef]
138. Wentzell, P.D.; Lohnes, M.T. Maximum likelihood principal component analysis with correlated measurement errors: Theoretical and practical considerations. *Chemom. Intell. Lab. Syst.* **1999**, *45*, 65–85. [CrossRef]

139. Wentzell, P.D.; Andrews, D.T.; Hamilton, D.C.; Faber, K.; Kowalski, B.R. Maximum likelihood principal component analysis. *J. Chemom.* **1997**, *11*, 339–366. [CrossRef]
140. Monteiro, A.d.R.D.; Feital, T.d.S.; Pinto, J.C. A Numerical Procedure for Multivariate Calibration Using Heteroscedastic Principal Components Regression. *Processes* **2021**, *9*, 1686. [CrossRef]
141. Narasimhan, S.; Bhatt, N. Deconstructing principal component analysis using a data reconciliation perspective. *Comput. Chem. Eng.* **2015**, *77*, 74–84. [CrossRef]
142. Ge, Z. Process Data Analytics via Probabilistic Latent Variable Models: A Tutorial Review. *Ind. Eng. Chem. Res.* **2018**, *57*, 12646–12661. [CrossRef]
143. Walczak, B.; Massart, D.L. Robust principal components regression as a detection tool for outliers. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 41–54. [CrossRef]
144. Xie, Y.L.; Wang, J.H.; Liang, Y.Z.; Sun, L.X.; Song, X.H.; Yu, R.Q. Robust principal component analysis by projection pursuit. *J. Chemom.* **1993**, *7*, 527–541. [CrossRef]
145. Wang, D.; Romagnoli, J. Robust multi-scale principal components analysis with applications to process monitoring. *J. Process Control* **2005**, *15*, 869–882. [CrossRef]
146. Jin, H.D.; Lee, Y.H.; Lee, G.; Han, C. Robust Recursive Principal Component Analysis Modeling for Adaptive Monitoring. *Ind. Eng. Chem. Res.* **2006**, *45*, 696–703. [CrossRef]
147. Tharrault, Y.; Mourot, G.; Ragot, J.; Maquin, D. Fault Detection and Isolation with Robust Principal Component Analysis. *Int. J. Appl. Math. Comput. Sci.* **2008**, *18*, 429–442. [CrossRef]
148. Lee, H.; Lee, J. Online update techniques for projection based Robust Principal Component Analysis. *ICT Express* **2015**, *1*, 59–62. [CrossRef]
149. Xiao, W.; Huang, X.; He, F.; Silva, J.; Emrani, S.; Chaudhuri, A. Online Robust Principal Component Analysis with Change Point Detection. *IEEE Trans. Multimed.* **2020**, *22*, 59–68. [CrossRef]
150. Zheng, J.; Zhu, J.; Chen, G.; Song, Z.; Ge, Z. Dynamic Bayesian network for robust latent variable modeling and fault classification. *Eng. Appl. Artif. Intell.* **2020**, *89*, 103475. [CrossRef]
151. Zhong, L.; Chang, Y.; Wang, F.; Gao, S. Distributed Missing Values Imputation Schemes for Plant-Wide Industrial Process Using Variational Bayesian Principal Component Analysis. *Ind. Eng. Chem. Res.* **2022**, *61*, 580–593. [CrossRef]
152. Bouwmans, T.; Zahzah, E.H. Robust PCA via Principal Component Pursuit: A review for a comparative evaluation in video surveillance. *Comput. Vis. Image Underst.* **2014**, *122*, 22–34. [CrossRef]
153. Severson, K.; Molaro, M.; Braatz, R. Principal Component Analysis of Process Datasets with Missing Values. *Processes* **2017**, *5*, 38. [CrossRef]
154. Zhu, J.; Ge, Z.; Song, Z.; Gao, F. Review and big data perspectives on robust data mining approaches for industrial process modeling with outliers and missing data. *Annu. Rev. Control* **2018**, *46*, 107–133. [CrossRef]
155. Hotelling, H. Relations Between Two Sets of Variates. *Biometrika* **1936**, *28*, 321–377. [CrossRef]
156. Wold, H. Estimation of Principal Components and Related Models by Iterative Least Squares. In *Multivariate Analysis: Proceedings of an International Symposium Held in Dayton, Ohio, 14–19 June 1965*; Academic Press: Minneapolis, MN, USA, 1966; pp. 391–420.
157. Gatius, F.; Miralbés, C.; David, C.; Puy, J. Comparison of CCA and PLS to explore and model NIR data. *Chemom. Intell. Lab. Syst.* **2017**, *164*, 76–82. [CrossRef]
158. Banner, M.; Alosert, H.; Spencer, C.; Cheeks, M.; Farid, S.S.; Thomas, M.; Goldrick, S. A decade in review: Use of data analytics within the biopharmaceutical sector. *Curr. Opin. Chem. Eng.* **2021**, *34*, 100758. [CrossRef] [PubMed]
159. Chen, Z. *Data-Driven Fault Detection for Industrial Processes*; Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2017. [CrossRef]
160. Zhang, K.; Peng, K.; Chu, R.; Dong, J. Implementing multivariate statistics-based process monitoring: A comparison of basic data modeling approaches. *Neurocomputing* **2018**, *290*, 172–184. [CrossRef]
161. Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* **2003**, *17*, 166–173. [CrossRef]
162. Abdi, H. Partial least squares regression and projection on latent structure regression (PLS Regression). *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 97–106. [CrossRef]
163. Rosipal, R.; Krämer, N. Overview and Recent Advances in Partial Least Squares. In *Subspace, Latent Structure and Feature Selection*; Saunders, C., Grobelnik, M., Gunn, S., Shawe-Taylor, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; Volume 3940, pp. 34–51. [CrossRef]
164. Ding, S.X. *Data-Driven Design of Fault Diagnosis and Fault-Tolerant Control Systems*; Advances in Industrial Control; Springer: London, UK, 2014. [CrossRef]
165. Severson, K.; Chaiwatanodom, P.; Braatz, R.D. Perspectives on process monitoring of industrial systems. *Annu. Rev. Control* **2016**, *42*, 190–200. [CrossRef]
166. Li, X.; Lin, T.; Mba, D. A Tutorial on Canonical Variate Analysis for Diagnosis and Prognosis. In *Smart Monitoring of Rotating Machinery for Industry 4.0*; Chaari, F., Chiementin, X., Zimroz, R., Bolaers, F., Haddar, M., Eds.; Applied Condition Monitoring; Springer International Publishing: Cham, Switerland, 2022; pp. 11–31. [CrossRef]
167. Gittins, R. *Canonical Analysis: A Review with Applications in Ecology*; Springer: Berlin/Heidelberg, Germany, 1985.
168. Ferrer, A.; Aguado, D.; Vidal-Puig, S.; Prats, J.M.; Zarzo, M. PLS: A versatile tool for industrial process improvement and optimization. *Appl. Stoch. Model. Bus. Ind.* **2008**, *24*, 551–567. [CrossRef]

169. Chen, J.; Liu, K.C. On-line batch process monitoring using dynamic PCA and dynamic PLS models. *Chem. Eng. Sci.* **2002**, *57*, 63–75. [CrossRef]
170. Kourti, T.; Nomikos, P.; MacGregor, J.F. Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS. *J. Process Control* **1995**, *5*, 277–284. [CrossRef]
171. Gregersen, L.; Jørgensen, S.B.; Andersen, M.Y. Industrial Fed-Batch Fermentation Monitoring. *IFAC Proc. Vol.* **1997**, *30*, 49–54. [CrossRef]
172. van Kollenburg, G.H.; van Es, J.; Gerretzen, J.; Lanters, H.; Bouman, R.; Koelewijn, W.; Davies, A.N.; Buydens, L.M.; van Manen, H.J.; Jansen, J.J. Understanding chemical production processes by using PLS path model parameters as soft sensors. *Comput. Chem. Eng.* **2020**, *139*, 106841. [CrossRef]
173. Agarwal, P.; Aghaee, M.; Tamer, M.; Budman, H. A novel unsupervised approach for batch process monitoring using deep learning. *Comput. Chem. Eng.* **2022**, *159*, 107694. [CrossRef]
174. Zhang, S.; Bao, X. Slow-varying batch process monitoring based on canonical variate analysis. *Can. J. Chem. Eng.* **2023**, *101*, 400–419. [CrossRef]
175. Qin, S.J. Recursive PLS algorithms for adaptive data modeling. *Comput. Chem. Eng.* **1998**, *22*, 503–514. [CrossRef]
176. Qin, S.; McAvoy, T. Nonlinear FIR modeling via a neural net PLS approach. *Comput. Chem. Eng.* **1996**, *20*, 147–159. [CrossRef]
177. Dong, Y.; Qin, S.J. Dynamic-Inner Partial Least Squares for Dynamic Data Modeling. *IFAC-PapersOnLine* **2015**, *48*, 117–122. [CrossRef]
178. Wang, X.; Kruger, U.; Lennox, B. Recursive partial least squares algorithms for monitoring complex industrial processes. *Control Eng. Pract.* **2003**, *11*, 613–632. [CrossRef]
179. Liu, Q.; Qin, S.J.; Chai, T. Quality-Relevant Monitoring and Diagnosis with Dynamic Concurrent Projection to Latent Structures. *IFAC Proc. Vol.* **2014**, *47*, 2740–2745. [CrossRef]
180. Lavoie, F.B.; Muteki, K.; Gosselin, R. Computing meaningful models of continuous data. *Chem. Eng. Res. Des.* **2019**, *148*, 321–335. [CrossRef]
181. Zhu, Q.; Joe Qin, S.; Dong, Y. Dynamic Latent Variable Regression for Inferential Sensor Modeling and Monitoring. *Comput. Chem. Eng.* **2020**, *137*, 106809. [CrossRef]
182. Yang, J.; Dong, J.; Shi, H.; Tan, S. Quality monitoring method based on enhanced canonical component analysis. *ISA Trans.* **2020**, *105*, 221–229. [CrossRef] [PubMed]
183. Lan, T.; Tong, C.; Shi, X.; Luo, L. Dynamic statistical process monitoring based on generalized canonical variate analysis. *J. Taiwan Inst. Chem. Eng.* **2020**, *112*, 78–86. [CrossRef]
184. Dong, Y.; Qin, S.J. Dynamic Inner Canonical Correlation Analysis based Process Monitoring. In Proceedings of the 2020 American Control Conference (ACC), Denver, CO, USA, 1–3 July 2020; pp. 3553–3558. [CrossRef]
185. Zhang, M.Q.; Luo, X.L. Modified canonical variate analysis based on dynamic kernel decomposition for dynamic nonlinear process quality monitoring. *ISA Trans.* **2021**, *108*, 106–120. [CrossRef] [PubMed]
186. Zhu, Q. Auto-regressive modeling with dynamic weighted canonical correlation analysis. *J. Process Control* **2020**, *95*, 32–44. [CrossRef]
187. Lou, S.; Wu, P.; Guo, L.; He, J.; Zhang, X.; Gao, J. Dynamic process monitoring using dynamic latent variable and canonical correlation analysis model. *Can. J. Chem. Eng.* **2020**, *99*, 986–998. [CrossRef]
188. Xu, B.; Zhu, Q. Online Quality Relevant Monitoring with Dynamic Weighted Partial Least Squares. *Ind. Eng. Chem. Res.* **2020**, *59*, 21124–21132. [CrossRef]
189. Qin, S.J.; Liu, Y.; Dong, Y. Plant-wide troubleshooting and diagnosis using dynamic embedded latent feature analysis. *Comput. Chem. Eng.* **2021**, *152*, 107392. [CrossRef]
190. Lou, S.; Wu, P.; Yang, C.; Xu, Y. Structured fault information-aided canonical variate analysis model for dynamic process monitoring. *J. Process Control* **2023**, *124*, 54–69. [CrossRef]
191. Zhang, T.; Yan, G.; Ren, M.; Cheng, L.; Li, R.; Xie, G. Dynamic transfer soft sensor for concept drift adaptation. *J. Process Control* **2023**, *123*, 50–63. [CrossRef]
192. Hu, X.; Wu, P.; Pan, H.; He, Y. A fault detection method based on sparse dynamic canonical correlation analysis. *Can. J. Chem. Eng.* 2023; *early view*. [CrossRef]
193. Zhang, Y.; Zhou, H.; Qin, S.J.; Chai, T. Decentralized Fault Diagnosis of Large-Scale Processes Using Multiblock Kernel Partial Least Squares. *IEEE Trans. Ind. Inform.* **2010**, *6*, 3–10. [CrossRef]
194. Peng, X.; Li, Z.; Zhong, W.; Qian, F.; Tian, Y. Concurrent Quality-Relevant Canonical Correlation Analysis for Nonlinear Continuous Process Decomposition and Monitoring. *Ind. Eng. Chem. Res.* **2020**, *59*, 8757–8768. [CrossRef]
195. Wu, P.; Lou, S.; Zhang, X.; He, J.; Gao, J. Novel Quality Relevant Process Monitoring based on Dynamic Locally Linear Embedding Concurrent Canonical Correlation Analysis. *Ind. Eng. Chem. Res.* **2020**, *59*, 21439–21457. [CrossRef]
196. Chen, Q.; Wang, Y. Key performance indicator related state monitoring based on kernel canonical correlation analysis. *Control Eng. Pract.* **2021**, *107*, 104692. [CrossRef]
197. Kong, X.; Luo, J.; Feng, X.; Liu, M. A General Quality-Related Nonlinear Process Monitoring Approach Based on Input–Output Kernel PLS. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–12. [CrossRef]
198. Chu, F.; Mo, S.; Hao, L.; Lu, N.; Wang, F. Operating Performance Assessment of Complex Nonlinear Industrial Process Based on Kernel Locally Linear Embedding PLS. *IEEE Trans. Autom. Sci. Eng.* **2022**, *21*, 593–605. [CrossRef]

199. Zhao, S.J.; Zhang, J.; Xu, Y.M. Performance monitoring of processes with multiple operating modes through multiple PLS models. *J. Process Control* **2006**, *16*, 763–772. [CrossRef]
200. Chen, Z.; Liu, C.; Ding, S.X.; Peng, T.; Yang, C.; Gui, W.; Shardt, Y.A.W. A Just In Time Learning Aided Canonical Correlation Analysis Method for Multimode Process Monitoring and Fault Detection. *IEEE Trans. Ind. Electron.* **2021**, *68*, 12. [CrossRef]
201. Sun, R.; Wang, Y. C IPLS IKPLS for Modeling and Detecting Nonlinear Multimode Processes. *Ind. Eng. Chem. Res.* **2021**, *60*, 1684–1698. [CrossRef]
202. Zhang, S.; Bao, X.; Wang, S. Common canonical variate analysis (CCVA) based modeling and monitoring for multimode processes. *Chem. Eng. Sci.* **2023**, *271*, 118581. [CrossRef]
203. Teppola, P.; Minkkinen, P. Wavelet-PLS regression models for both exploratory data analysis and process monitoring. *J. Chemom.* **2000**, *14*, 383–399. [CrossRef]
204. Lee, H.W.; Lee, M.W.; Park, J.M. Multi-scale extension of PLS algorithm for advanced on-line process monitoring. *Chemom. Intell. Lab. Syst.* **2009**, *98*, 201–212. [CrossRef]
205. Botre, C.; Mansouri, M.; Karim, M.N.; Nounou, H.; Nounou, M. Multiscale PLS-based GLRT for fault detection of chemical processes. *J. Loss Prev. Process Ind.* **2017**, *46*, 143–153. [CrossRef]
206. MacGregor, J.F.; Jaeckle, C.; Kiparissides, C.; Koutoudi, M. Process monitoring and diagnosis by multiblock PLS methods. *AIChE J.* **1994**, *40*, 826–838. [CrossRef]
207. Choi, S.W.; Lee, I.B. Multiblock PLS-based localized process diagnosis. *J. Process Control* **2005**, *15*, 295–306. [CrossRef]
208. Sun, R. Multiblock global orthogonal projections to latent structures for fault diagnosis. *Chemom. Intell. Lab. Syst.* **2020**, *204*, 104092. [CrossRef]
209. Song, B.; Shi, H.; Tan, S.; Tao, Y. Multi-Subspace Orthogonal Canonical Correlation Analysis for Quality Related Plant Wide Process Monitoring. *IEEE Trans. Ind. Inform.* **2020**, *17*, 6368–6378. [CrossRef]
210. Jiang, Q.; Yan, X. Neighborhood Stable Correlation Analysis for Robust Monitoring of Multiunit Chemical Processes. *Ind. Eng. Chem. Res.* **2020**, *59*, 16695–16707. [CrossRef]
211. Chen, X.; Wang, J.; Ding, S.X. Complex System Monitoring Based on Distributed Least Squares Method. *IEEE Trans. Autom. Sci. Eng.* **2020**, *18*, 1892–1900. [CrossRef]
212. Zhu, Q.X.; Wang, X.W.; Li, K.; Xu, Y.; He, Y.L. Enhanced multicorrelation block process monitoring and abnormity root cause analysis for distributed industrial process: A visual data-driven approach. *J. Process Control* **2022**, *118*, 1–15. [CrossRef]
213. Fuente, M.J.D.L.; Sainz-Palmero, G.I.; Galende-Hern¡ndez, M. Dynamic Decentralized Monitoring for Large-Scale Industrial Processes Using Multiblock Canonical Variate Analysis Based Regression. *IEEE Access* **2023**, *11*, 26611–26623. [CrossRef]
214. Zheng, J.; Song, Z.; Ge, Z. Probabilistic learning of partial least squares regression model: Theory and industrial applications. *Chemom. Intell. Lab. Syst.* **2016**, *158*, 80–90. [CrossRef]
215. Li, Q.; Pan, F.; Zhao, Z. Concurrent probabilistic PLS regression model and its applications in process monitoring. *Chemom. Intell. Lab. Syst.* **2017**, *171*, 40–54. [CrossRef]
216. Li, Q.; Pan, F.; Zhao, Z.; Yu, J. Process Modeling and Monitoring With Incomplete Data Based on Robust Probabilistic Partial Least Square Method. *IEEE Access* **2018**, *6*, 10160–10168. [CrossRef]
217. Qin, Y.; Lou, Z.; Wang, Y.; Lu, S.; Sun, P. An analytical partial least squares method for process monitoring. *Control Eng. Pract.* **2022**, *124*, 105182. [CrossRef]
218. Lee, H.W.; Lee, M.W.; Park, J.M. Robust Adaptive Partial Least Squares Modeling of a Full-Scale Industrial Wastewater Treatment Process. *Ind. Eng. Chem. Res.* **2007**, *46*, 955–964. [CrossRef]
219. Yi, H.; Hehe, M.; Hongbo, S. Fault detection for chemical process based on robust PLS. In Proceedings of the 2013 25th Chinese Control and Decision Conference (CCDC), Guiyang, China, 25–27 May 2013; pp. 4947–4952. [CrossRef]
220. Yin, S.; Wang, G.; Yang, X. Robust PLS approach for KPI-related prediction and diagnosis against outliers and missing data. *Int. J. Syst. Sci.* **2014**, *45*, 1375–1382. [CrossRef]
221. Cheng, H.; Wu, J.; Huang, D.; Liu, Y.; Wang, Q. Robust adaptive boosted canonical correlation analysis for quality-relevant process monitoring of wastewater treatment. *ISA Trans.* **2021**, *117*, 210–220. [CrossRef] [PubMed]
222. Luo, L.; Wang, W.; Bao, S.; Peng, X.; Peng, Y. Robust and sparse canonical correlation analysis for fault detection and diagnosis using training data with outliers. *Expert Syst. Appl.* **2024**, *236*, 121434. [CrossRef]
223. Li, Z.; Yu, Y.; Pan, X.; Karim, M.N. Effect of dataset size on modeling and monitoring of chemical processes. *Chem. Eng. Sci.* **2020**, *227*, 115928. [CrossRef]
224. Yang, X.; Liu, W.; Liu, W.; Tao, D. A Survey on Canonical Correlation Analysis. *IEEE Trans. Knowl. Data Eng.* **2021**, *33*, 2349–2368. [CrossRef]
225. Fisher, R.A. The Use of Multiple Measurements in Taxonomic Problems. *Ann. Eugen.* **1936**, *7*, 179–188. [CrossRef]
226. He, Q.P.; Qin, S.J.; Wang, J. A new fault diagnosis method using fault directions in Fisher discriminant analysis. *AIChE J.* **2005**, *51*, 555–571. [CrossRef]
227. Cho, H.W. Identification of contributing variables using kernel-based discriminant modeling and reconstruction. *Expert Syst. Appl.* **2007**, *33*, 274–285. [CrossRef]
228. He, X.B.; Wang, W.; Yang, Y.P.; Yang, Y.H. Variable-weighted Fisher discriminant analysis for process fault diagnosis. *J. Process Control* **2009**, *19*, 923–931. [CrossRef]

229. Wang, N.; Yuan, Z.; Wang, D. Improving Process Fault Detection and Diagnosis Using Robust PCA and Robust FDA. In Proceedings of the 2009 WRI World Congress on Computer Science and Information Engineering, Los Angeles, CA, USA, 31 March–2 April 2009; pp. 54–59. [CrossRef]
230. Zhong, S.; Wen, Q.; Ge, Z. Semi-supervised Fisher discriminant analysis model for fault classification in industrial processes. *Chemom. Intell. Lab. Syst.* **2014**, *138*, 203–211. [CrossRef]
231. Jiang, B.; Zhu, X.; Huang, D.; Paulson, J.A.; Braatz, R.D. A combined canonical variate analysis and Fisher discriminant analysis (CVA–FDA) approach for fault diagnosis. *Comput. Chem. Eng.* **2015**, *77*, 1–9. [CrossRef]
232. Adil, M.; Abid, M.; Khan, A.; Mustafa, G.; Ahmed, N. Exponential discriminant analysis for fault diagnosis. *Neurocomputing* **2016**, *171*, 1344–1353. [CrossRef]
233. Deng, X.; Tian, X.; Chen, S.; Harris, C.J. Statistics local fisher discriminant analysis for industrial process fault classification. In Proceedings of the 2016 UKACC 11th International Conference on Control (CONTROL), Belfast, UK, 31 August–2 September 2016. [CrossRef]
234. Feng, J.; Wang, J.; Zhang, H.; Han, Z. Fault Diagnosis Method of Joint Fisher Discriminant Analysis Based on the Local and Global Manifold Learning and Its Kernel Version. *IEEE Trans. Autom. Sci. Eng.* **2016**, *13*, 122–133. [CrossRef]
235. Md Nor, N.; Hussain, M.A.; Hassan, C.R.C. Process Monitoring and Fault Detection in Non-Linear Chemical Process Based On Multi-Scale Kernel Fisher Discriminant Analysis. In *Computer Aided Chemical Engineering*; Elsevier: Amsterdam, The Netherlands, 2015; Volume 37, pp. 1823–1828. [CrossRef]
236. Md Nor, N.; Hussain, M.A.; Che Hassan, C.R. Fault diagnosis and classification framework using multi-scale classification based on kernel Fisher discriminant analysis for chemical process system. *Appl. Soft Comput.* **2017**, *61*, 959–972. [CrossRef]
237. Jia, Z.; Zeng, Y.; Zhang, Y.; Liang, W. Local class-specific discriminant analysis with variable weighting and its application in fault diagnosis. *Comput. Chem. Eng.* **2020**, *141*, 107023. [CrossRef]
238. Shang, C.; Zhao, L.; Huang, X.; Ye, H.; Huang, D. Group-sparsity-enforcing fault discrimination and estimation with dynamic process data. *J. Process Control* **2021**, *105*, 236–249. [CrossRef]
239. Lou, C.; Atoui, M.A.; Li, X. Novel online discriminant analysis based schemes to deal with observations from known and new classes: Application to industrial systems. *Eng. Appl. Artif. Intell.* **2022**, *111*, 104811. [CrossRef]
240. Ding, Z.; Xu, Y.; Zhong, K. Exponential Local Fisher Discriminant Analysis with Sparse Variables Selection: A Novel Fault Diagnosis Scheme for Industry Application. *Machines* **2023**, *11*, 1066. [CrossRef]
241. Jutten, C.; Herault, J. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Process.* **1991**, *24*, 1–10. [CrossRef]
242. Zhang, S.; Wang, F.; Wang, S.; Tan, S.; Chang, Y. Multivariate process monitoring based on the distribution test of the data. In Proceedings of the Proceeding of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, 29 June–4 July 2014; pp. 2723–2728. [CrossRef]
243. Kano, M.; Hasebe, S.; Hashimoto, I.; Ohno, H. Evolution of multivariate statistical process control: Application of independent component analysis and external analysis. *Comput. Chem. Eng.* **2004**, *28*, 1157–1166. [CrossRef]
244. Lee, J.M.; Yoo, C.; Lee, I.B. Statistical process monitoring with independent component analysis. *J. Process Control* **2004**, *14*, 467–485. [CrossRef]
245. Lee, J.M.; Yoo, C.; Lee, I.B. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chem. Eng. Sci.* **2004**, *59*, 2995–3006. [CrossRef]
246. Albazzaz, H.; Wang, X.Z. Statistical Process Control Charts for Batch Operations Based on Independent Component Analysis. *Ind. Eng. Chem. Res.* **2004**, *43*, 6731–6741. [CrossRef]
247. Kano, M.; Tanaka, S.; Hasebe, S.; Hashimoto, I.; Ohno, H. Monitoring independent components for fault detection. *AIChE J.* **2003**, *49*, 969–976. [CrossRef]
248. Lee, J.M.; Yoo, C.K.; Lee, I. New monitoring technique with an ICA algorithm in the wastewater treatment process. *Water Sci. Technol.* **2003**, *47*, 49–56. [CrossRef] [PubMed]
249. Chen, J.; Yu, J. Independent Component Analysis Mixture Model Based Dissimilarity Method for Performance Monitoring of Non-Gaussian Dynamic Processes with Shifting Operating Conditions. *Ind. Eng. Chem. Res.* **2014**, *53*, 5055–5066. [CrossRef]
250. Hsu, C.C.; Chen, M.C.; Chen, L.S. Integrating independent component analysis and support vector machine for multivariate process monitoring. *Comput. Ind. Eng.* **2010**, *59*, 145–156. [CrossRef]
251. Hsu, C.C.; Chen, M.C.; Chen, L.S. Intelligent ICA–SVM fault detector for non-Gaussian multivariate process monitoring. *Expert Syst. Appl.* **2010**, *37*, 3264–3273. [CrossRef]
252. Hsu, C.C.; Chen, M.C.; Chen, L.S. A novel process monitoring approach with dynamic independent component analysis. *Control Eng. Pract.* **2010**, *18*, 242–253. [CrossRef]
253. Xu, Y.; Deng, X. Fault detection of multimode non-Gaussian dynamic process using dynamic Bayesian independent component analysis. *Neurocomputing* **2016**, *200*, 70–79. [CrossRef]
254. Zhu, Q.X.; Meng, Q.Q.; He, Y.L. Novel Multidimensional Feature Pattern Classification Method and Its Application to Fault Diagnosis. *Ind. Eng. Chem. Res.* **2017**, *56*, 8906–8916. [CrossRef]
255. Li, S.; Zhou, X.; Pan, F.; Shi, H.; Li, K.; Wang, Z. Correlated and weakly correlated fault detection based on variable division and ICA. *Comput. Ind. Eng.* **2017**, *112*, 320–335. [CrossRef]

256. Lughofer, E.; Zavoianu, A.C.; Pollak, R.; Pratama, M.; Meyer-Heye, P.; Zörrer, H.; Eitzinger, C.; Radauer, T. On-line anomaly detection with advanced independent component analysis of multi-variate residual signals from causal relation networks. *Inf. Sci.* **2020**, *537*, 425–451. [CrossRef]
257. Xu, C.; Huang, D.; Li, D.; Liu, Y. Novel Process Monitoring Approach Enhanced by a Complex Independent Component Analysis Algorithm with Applications for Wastewater Treatment. *Ind. Eng. Chem. Res.* **2021**, *60*, 13914–13926. [CrossRef]
258. Xu, C.; Huang, D.; Cai, B.; Chen, H.; Liu, Y. A complex-valued slow independent component analysis based incipient fault detection and diagnosis method with applications to wastewater treatment processes. *ISA Trans.* **2023**, *135*, 213–232. [CrossRef] [PubMed]
259. Liu, M.; Kong, X.; Luo, J.; Yang, L. Fault detection and diagnosis in a non-Gaussian process with modified kernel independent component regression. *Can. J. Chem. Eng.* **2023**, *102*, 781–802. [CrossRef]
260. Zhang, S.; Wang, F.; Zhao, L.; Wang, S.; Chang, Y. A Novel Strategy of the Data Characteristics Test for Selecting a Process Monitoring Method Automatically. *Ind. Eng. Chem. Res.* **2016**, *55*, 1642–1654. [CrossRef]
261. Tharwat, A. Independent component analysis: An introduction. *Appl. Comput. Inform.* **2021**, *17*, 222–249. [CrossRef]
262. Comon, P. Independent component analysis, A new concept? *Signal Process.* **1994**, *36*, 287–314. [CrossRef]
263. Palla, G.P.L.; Pani, A.K. Independent component analysis application for fault detection in process industries: Literature review and an application case study for fault detection in multiphase flow systems. *Measurement* **2023**, *209*, 112504. [CrossRef]
264. Wiskott, L.; Sejnowski, T.J. Slow Feature Analysis: Unsupervised Learning of Invariances. *Neural Comput.* **2002**, *14*, 715–770. [CrossRef]
265. Gao, X.; Shardt, Y.A.W. EVOLVE·INFOMAX: An Unsupervised Learning Principle of Invariances for Nonlinear Dynamic Systems. *Ind. Eng. Chem. Res.* **2022**, *61*, 18004–18016. [CrossRef]
266. Li, Q.; Zhao, Z.; Wang, L. Slowness or Autocorrelation? A serial correlation feature analysis method and its application in process monitoring. *J. Process Control* **2023**, *121*, 1–12. [CrossRef]
267. Liu, J.; Mu, G.; Chen, J. Tensor slow feature analysis and its applications for batch process monitoring. *Comput. Chem. Eng.* **2023**, *173*, 108207. [CrossRef]
268. Zhang, X.; Ma, L.; Peng, K.; Zhang, C. A quality-related distributed fault detection method for large-scale sequential processes. *Control Eng. Pract.* **2022**, *127*, 105308. [CrossRef]
269. Saafan, H.; Zhu, Q. Improved manifold sparse slow feature analysis for process monitoring. *Comput. Chem. Eng.* **2022**, *164*, 107905. [CrossRef]
270. Zhang, C.; Dong, J.; Peng, K.; You, P. Dynamic industrial process monitoring based on concurrent fast and slow-time-varying feature analytics. *Can. J. Chem. Eng.* **2022**, *100*, S141–S153. [CrossRef]
271. Li, S.; Luo, J.; Hu, Y. Toward Interpretable Process Monitoring: Slow Feature Analysis-Aided Autoencoder for Spatiotemporal Process Feature Learning. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–11. [CrossRef]
272. Xu, Y.; Jia, M.; Mao, Z. A novel auto-regressive dynamic slow feature analysis method for dynamic chemical process monitoring. *Chem. Eng. Sci.* **2022**, *248*, 117236. [CrossRef]
273. Xu, Y.; Jia, M.; Mao, Z.; Li, H. A novel kernel dynamic inner slow feature analysis method for dynamic nonlinear process concurrent monitoring of operating point deviations and process dynamics anomalies. *J. Process Control* **2022**, *110*, 59–75. [CrossRef]
274. Li, C.; Zhou, Z.; Wen, C.; Li, Z. Fault Detection of Non-Gaussian and Nonlinear Processes Based on Independent Slow Feature Analysis. *ACS Omega* **2022**, *7*, 6978–6990. [CrossRef]
275. Gao, X.; Shardt, Y.A.W. Dynamic system modelling and process monitoring based on long-term dependency slow feature analysis. *J. Process Control* **2021**, *105*, 27–47. [CrossRef]
276. Puli, V.K.; Raveendran, R.; Huang, B. Complex probabilistic slow feature extraction with applications in process data analytics. *Comput. Chem. Eng.* **2021**, *154*, 107456. [CrossRef]
277. Zhong, K.; Ma, D.; Han, M. Distributed dynamic process monitoring based on dynamic slow feature analysis with minimal redundancy maximal relevance. *Control Eng. Pract.* **2020**, *104*, 104627. [CrossRef]
278. Huang, J.; Yang, X.; Yan, X. Slow feature analysis-independent component analysis based integrated monitoring approach for industrial processes incorporating dynamic and static characteristics. *Control Eng. Pract.* **2020**, *102*, 104558. [CrossRef]
279. Xu, X.; Ding, J. Decentralized dynamic process monitoring based on manifold regularized slow feature analysis. *J. Process Control* **2021**, *98*, 79–91. [CrossRef]
280. Zhang, H.; Deng, X.; Zhang, Y.; Hou, C.; Li, C. Dynamic nonlinear batch process fault detection and identification based on two-directional dynamic kernel slow feature analysis. *Can. J. Chem. Eng.* **2021**, *99*, 306–333. [CrossRef]
281. Wang, J.; Zhao, Z.; Liu, F. Robust Slow Feature Analysis for Statistical Process Monitoring. *Ind. Eng. Chem. Res.* **2020**, *59*, 12504–12513. [CrossRef]
282. Li, C.; Wen, C.; Zhou, Z. An industrial process fault diagnosis method based on independent slow feature analysis and stacked sparse autoencoder network. *J. Frankl. Inst.* **2024**, *361*, 234–247. [CrossRef]
283. Song, P.; Zhao, C. Slow Down to Go Better: A Survey on Slow Feature Analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–21. [CrossRef]
284. Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2009.

285. Venkatasubramanian, V. The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE J.* **2019**, *65*, 466–478. [CrossRef]
286. Pan, Y. Heading toward Artificial Intelligence 2.0. *Engineering* **2016**, *2*, 409–413. [CrossRef]
287. Kaplan, A.; Haenlein, M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Bus. Horizons* **2019**, *62*, 15–25. [CrossRef]
288. Lee, J.H.; Shin, J.; Realff, M.J. Machine learning: Overview of the recent progresses and implications for the process systems engineering field. *Comput. Chem. Eng.* **2018**, *114*, 111–121. [CrossRef]
289. Schweidtmann, A.M.; Esche, E.; Fischer, A.; Kloft, M.; Repke, J.; Sager, S.; Mitsos, A. Machine Learning in Chemical Engineering: A Perspective. *Chem. Ing. Tech.* **2021**, *93*, 2029–2039. [CrossRef]
290. Dobbelaere, M.R.; Plehiers, P.P.; Van de Vijver, R.; Stevens, C.V.; Van Geem, K.M. Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats. *Engineering* **2021**, *7*, 1201–1211. [CrossRef]
291. Mowbray, M.; Vallerio, M.; Perez-Galvan, C.; Zhang, D.; Del Rio Chanona, A.; Navarro-Brull, F.J. Industrial data science – a review of machine learning applications for chemical and process industries. *React. Chem. Eng.* **2022**, *7*, 1471–1509. [CrossRef]
292. Gao, H.; Zhu, L.T.; Luo, Z.H.; Fraga, M.A.; Hsing, I.M. Machine Learning and Data Science in Chemical Engineering. *Ind. Eng. Chem. Res.* **2022**, *61*, 8357–8358. [CrossRef]
293. Fuentes-Cortés, L.F.; Flores-Tlacuahuac, A.; Nigam, K.D.P. Machine Learning Algorithms Used in PSE Environments: A Didactic Approach and Critical Perspective. *Ind. Eng. Chem. Res.* **2022**, *61*, 8932–8962. [CrossRef]
294. Khan, N.; Ammar Taqvi, S.A. Machine Learning an Intelligent Approach in Process Industries: A Perspective and Overview. *ChemBioEng Rev.* **2023**, *10*, 195–221. [CrossRef]
295. Zavala, V.M. Outlook: How I Learned to Love Machine Learning (A Personal Perspective on Machine Learning in Process Systems Engineering). *Ind. Eng. Chem. Res.* **2023**, *62*, 8995–9005. [CrossRef]
296. Mayo, M. The Essence of Machine Learning. 2018. Available online: https://www.kdnuggets.com/2018/12/essence-machine-learning.html (accessed on 16 January 2024).
297. Mitchell, T.M. *Machine Learning*, 1st ed.; McGraw-Hill Science/Engineering/Math: New York, NY, USA, 1997.
298. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, UK, 2016.
299. Li, S.; Luo, J.; Hu, Y. Semi-supervised process fault classification based on convolutional ladder network with local and global feature fusion. *Comput. Chem. Eng.* **2020**, *140*, 106843. [CrossRef]
300. Wei, C.; Song, Z. Generalized Semisupervised Self-Optimizing Kernel Model for Quality-Related Industrial Process Monitoring. *IEEE Trans. Ind. Electron.* **2020**, *67*, 10876–10886. [CrossRef]
301. Zhang, Y.; Yu, K.; Lei, Z.; Ge, J.; Xu, Y.; Li, Z.; Ren, Z.; Feng, K. Integrated intelligent fault diagnosis approach of offshore wind turbine bearing based on information stream fusion and semi-supervised learning. *Expert Syst. Appl.* **2023**, *232*, 120854. [CrossRef]
302. Shin, J.; Badgwell, T.A.; Liu, K.H.; Lee, J.H. Reinforcement Learning—Overview of recent progress and implications for process control. *Comput. Chem. Eng.* **2019**, *127*, 282–294. [CrossRef]
303. Torres, A.d.R.; Andreiana, D.S.; Roldán, A.O.; Bustos, A.H.; Galicia, L.E.A. A Review of Deep Reinforcement Learning Approaches for Smart Manufacturing in Industry 4.0 and 5.0 Framework. *Appl. Sci.* **2022**, *12*, 12377. [CrossRef]
304. Faria, R.d.R.; Capron, B.D.O.; Secchi, A.R.; de Souza, M.B. Where Reinforcement Learning Meets Process Control: Review and Guidelines. *Processes* **2022**, *10*, 2311. [CrossRef]
305. Nian, R.; Liu, J.; Huang, B. A review On reinforcement learning: Introduction and applications in industrial process control. *Comput. Chem. Eng.* **2020**, *139*, 106886. [CrossRef]
306. Zhang, D.; Fu, Y.; Lin, Z.; Gao, Z. A reinforcement learning based fault diagnosis for autoregressive-moving-average model. In Proceedings of the IECON 2017—43rd Annual Conference of the IEEE Industrial Electronics Society, Beijing, China, 29 October–1 November 2017; pp. 7067–7072. [CrossRef]
307. Acernese, A.; Yerudkar, A.; Del Vecchio, C. A Novel Reinforcement Learning-based Unsupervised Fault Detection for Industrial Manufacturing Systems. In Proceedings of the 2022 American Control Conference (ACC), Atlanta, GA, USA, 8–10 June 2022; pp. 2650–2655. [CrossRef]
308. Chen, S.; Jiang, Q. Distributed Robust Process Monitoring Based on Optimized Denoising Autoencoder With Reinforcement Learning. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–11. [CrossRef]
309. Venkatasubramanian, V.; Chan, K. A neural network methodology for process fault diagnosis. *AIChE J.* **1989**, *35*, 1993–2002. [CrossRef]
310. Hoskins, J.C.; Kaliyur, K.M.; Himmelblau, D.M. Fault diagnosis in complex chemical plants using artificial neural networks. *AIChE J.* **1991**, *37*, 137–141. [CrossRef]
311. Kramer, M. Autoassociative neural networks. *Comput. Chem. Eng.* **1992**, *16*, 313–328. [CrossRef]
312. Jack, L.B.; Nandi, A.K. Support vector machines for detection and characterization of rolling element bearing faults. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2001**, *215*, 1065–1074. [CrossRef]
313. Venkatasubramanian, V.; Rengaswamy, R.; Ka, S.N.; Yin, K. A review of process fault detection and diagnosis Part III: Process history based methods. *Comput. Chem. Eng.* **2003**, *27*, 327–346. [CrossRef]
314. Aldrich, C.; Auret, L. *Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods*; Springer: London, UK, 2013. [CrossRef]

315. Aggarwal, C.C. *Neural Networks and Deep Learning: A Textbook*; Springer International Publishing: Cham, Switerland, 2018. [CrossRef]
316. Chowdhary, K.R. Statistical Learning Theory. In *Fundamentals of Artificial Intelligence*; Chowdhary, K., Ed.; Springer: New Delhi, India, 2020; pp. 415–443. [CrossRef]
317. Nielsen, M. *Neural Networks and Deep Learning*; Determination Press: San Francisco, CA, USA, 2015.
318. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [CrossRef]
319. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
320. Werbos, P.J. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. Thesis, Harvard University, Cambridge, MA, USA, 1974.
321. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed.; O'Reilly Media: Sebastopol, CA, USA, 2022.
322. Hinton, G.E.; Osindero, S.; Teh, Y.W. A Fast Learning Algorithm for Deep Belief Nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef]
323. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13 May 2010; pp. 249–256.
324. Chen, J.X. The Evolution of Computing: AlphaGo. *Comput. Sci. Eng.* **2016**, *18*, 4–7. [CrossRef]
325. Wang, F.; Zhang, J.J.; Zheng, X.; Wang, X.; Yuan, Y.; Dai, X.; Zhang, J.; Yang, L. Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond. *IEEE/CAA J. Autom. Sin.* **2016**, *3*, 113–120.
326. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving Language Understanding by Generative Pre-Training. 2018. Available online: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf (accessed on 16 January 2024).
327. Sivaram, A.; Das, L.; Venkatasubramanian, V. Hidden representations in deep neural networks: Part 1. Classification problems. *Comput. Chem. Eng.* **2020**, *134*, 106669. [CrossRef]
328. Das, L.; Sivaram, A.; Venkatasubramanian, V. Hidden representations in deep neural networks: Part 2. Regression problems. *Comput. Chem. Eng.* **2020**, *139*, 106895. [CrossRef]
329. Sengupta, S.; Basak, S.; Saikia, P.; Paul, S.; Tsalavoutis, V.; Atiah, F.; Ravi, V.; Peters, A. A review of deep learning with special emphasis on architectures, applications and recent trends. *Knowl.-Based Syst.* **2020**, *194*, 105596. [CrossRef]
330. Zhou, J.; Zhu, Y. A critical look at deep neural network for dynamic system modeling. *arXiv* **2023**, arXiv:2301.11604.
331. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]
332. Pirdashti, M.; Curteanu, S.; Kamangar, M.H.; Hassim, M.H.; Khatami, M.A. Artificial neural networks: Applications in chemical engineering. *Rev. Chem. Eng.* **2013**, *29*, 205–239. [CrossRef]
333. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* **2016**, arXiv:1409.0473. [CrossRef]
334. Cheng, J.; Dong, L.; Lapata, M. Long Short-Term Memory-Networks for Machine Reading. *arXiv* **2016**, arXiv:1601.06733. [CrossRef]
335. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [CrossRef]
336. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *arXiv* **2017**, arXiv:1706.03762. [CrossRef]
337. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385.
338. Charte, D.; Charte, F.; del Jesus, M.J.; Herrera, F. An analysis on the use of autoencoders for representation learning: Fundamentals, learning task case studies, explainability and challenges. *Neurocomputing* **2020**, *404*, 93–107. [CrossRef]
339. Yang, Z.; Ge, Z. Monitoring and prediction of big process data with deep latent variable models and parallel computing. *J. Process Control* **2020**, *92*, 19–34. [CrossRef]
340. Qian, J.; Song, Z.; Yao, Y.; Zhu, Z.; Zhang, X. A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes. *Chemom. Intell. Lab. Syst.* **2022**, *231*, 104711. [CrossRef]
341. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *arXiv* **2014**, arXiv:1406.2661.
342. Brophy, E.; Wang, Z.; She, Q.; Ward, T. Generative Adversarial Networks in Time Series: A Systematic Literature Review. *ACM Comput. Surv.* **2023**, *55*, 199:1–199:31. [CrossRef]
343. Sabuhi, M.; Zhou, M.; Bezemer, C.P.; Musilek, P. Applications of Generative Adversarial Networks in Anomaly Detection: A Systematic Literature Review. *IEEE Access* **2021**, *9*, 161003–161029. [CrossRef]
344. Fischer, A.; Igel, C. An Introduction to Restricted Boltzmann Machines. In Proceedings of the Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Buenos Aires, Argentina, 3–6 September 2012; Lecture Notes in Computer Science; Alvarez, L., Mejail, M., Gomez, L., Jacobo, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 14–36. [CrossRef]
345. Yu, J.; Yan, X. Active features extracted by deep belief network for process monitoring. *ISA Trans.* **2019**, *84*, 247–261. [CrossRef] [PubMed]
346. Kohonen, T. The self-organizing map. *Proc. IEEE* **1990**, *78*, 1464–1480. [CrossRef]

347. Yin, H. The Self-Organizing Maps: Background, Theories, Extensions and Applications. In *Computational Intelligence: A Compendium*; Fulcher, J., Jain, L.C., Eds.; Studies in Computational Intelligence; Springer: Berlin/Heidelberg, Germany, 2008; pp. 715–762. [CrossRef]
348. Ahmad, A.; Hamid, M.K.A. Neural Networks for Process Monitoring, Control and Fault Detection: Application to Tennessee Eastman Plant. In Proceedings of the Malaysian Science and Technology Congress, Melaka, Malaysia, 8–10 October 2001.
349. Wu, H.; Zhao, J. Deep convolutional neural network model based chemical process fault diagnosis. *Comput. Chem. Eng.* **2018**, *115*, 185–197. [CrossRef]
350. Zhao, H.; Lai, Z. Neighborhood preserving neural network for fault detection. *Neural Netw.* **2019**, *109*, 6–18. [CrossRef]
351. Xavier, G.M.; de Seixas, J.M. Fault Detection and Diagnosis in a Chemical Process using Long Short-Term Memory Recurrent Neural Network. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018. [CrossRef]
352. Zhao, H.; Sun, S.; Jin, B. Sequential Fault Diagnosis Based on LSTM Neural Network. *IEEE Access* **2018**, *6*, 12929–12939. [CrossRef]
353. Yuan, J.; Tian, Y. An Intelligent Fault Diagnosis Method Using GRU Neural Network towards Sequential Data in Dynamic Processes. *Processes* **2019**, *7*, 152. [CrossRef]
354. Shahnazari, H. Fault diagnosis of nonlinear systems using recurrent neural networks. *Chem. Eng. Res. Des.* **2020**, *153*, 233–245. [CrossRef]
355. Ouyang, H.; Zeng, J.; Li, Y.; Luo, S. Fault Detection and Identification of Blast Furnace Ironmaking Process Using the Gated Recurrent Unit Network. *Processes* **2020**, *8*, 391. [CrossRef]
356. Han, Y.; Ding, N.; Geng, Z.; Wang, Z.; Chu, C. An optimized long short-term memory network based fault diagnosis model for chemical processes. *J. Process Control* **2020**, *92*, 161–168. [CrossRef]
357. Peng, P.; Zhang, W.; Zhang, Y.; Xu, Y.; Wang, H.; Zhang, H. Cost sensitive active learning using bidirectional gated recurrent neural networks for imbalanced fault diagnosis. *Neurocomputing* **2020**, *407*, 232–245. [CrossRef]
358. Kang, J.L. Visualization analysis for fault diagnosis in chemical processes using recurrent neural networks. *J. Taiwan Inst. Chem. Eng.* **2020**, *112*, 137–151. [CrossRef]
359. Agarwal, P.; Gonzalez, J.I.M.; Elkamel, A.; Budman, H. Hierarchical Deep Recurrent Neural Network based Method for Fault Detection and Diagnosis. *arXiv* **2020**, arXiv:2012.03861.
360. Lemos, T.; Campos, L.F.; Melo, A.; Clavijo, N.; Soares, R.; Câmara, M.; Feital, T.; Anzai, T.; Pinto, J.C. Echo State network based soft sensor for Monitoring and Fault Detection of Industrial Processes. *Comput. Chem. Eng.* **2021**, *155*, 107512. [CrossRef]
361. Xiong, S.; Zhou, L.; Dai, Y.; Ji, X. Attention-based LSTM fully convolutional network for chemical process fault diagnosis. *Chin. J. Chem. Eng.* **2022**, *56*, 1–14. [CrossRef]
362. Khalifa, R.M.; Yacout, S.; Bassetto, S. Root cause analysis of an out-of-control process using a logical analysis of data regression model and exponential weighted moving average. *J. Intell. Manuf.* **2023**. [CrossRef]
363. Agarwal, P.; Gonzalez, J.I.M.; Elkamel, A.; Budman, H. Hierarchical Deep LSTM for Fault Detection and Diagnosis for a Chemical Process. *Processes* **2022**, *10*, 2557. [CrossRef]
364. Ji, C.; Ma, F.; Wang, J.; Sun, W. Profitability related industrial-scale batch processes monitoring via deep learning based soft sensor development. *Comput. Chem. Eng.* **2023**, *170*, 108125. [CrossRef]
365. Deng, W.; Li, Y.; Huang, K.; Wu, D.; Yang, C.; Gui, W. LSTMED: An uneven dynamic process monitoring method based on LSTM and Autoencoder neural network. *Neural Netw.* **2023**, *158*, 30–41. [CrossRef] [PubMed]
366. Li, J.; Qin, S.J. Applying and dissecting LSTM neural networks and regularized learning for dynamic inferential modeling. *Comput. Chem. Eng.* **2023**, *175*, 108264. [CrossRef]
367. Liu, S.; Liu, Q.; Ahmed, S.; Wang, J.; Lei, F.; Zhao, D. A method for the early prediction of abnormal conditions in chemical processes combined with physical knowledge and the data-driven model. *J. Loss Prev. Process Ind.* **2023**, *86*, 105185. [CrossRef]
368. Zheng, S.; Zhao, J. A new unsupervised data mining method based on the stacked autoencoder for chemical process fault diagnosis. *Comput. Chem. Eng.* **2020**, *135*, 106755. [CrossRef]
369. Gao, X.; Yang, F.; Feng, E. A process fault diagnosis method using multi-time scale dynamic feature extraction based on convolutional neural network. *Can. J. Chem. Eng.* **2020**, *98*, 1280–1292. [CrossRef]
370. Yu, J.; Zhang, C.; Wang, S. Multichannel one-dimensional convolutional neural network-based feature learning for fault diagnosis of industrial processes. *Neural Comput. Appl.* **2020**, *33*, 3085–3104. [CrossRef]
371. Olivier, J.; Aldrich, C. Dynamic Monitoring of Grinding Circuits by Use of Global Recurrence Plots and Convolutional Neural Networks. *Minerals* **2020**, *10*, 958. [CrossRef]
372. Yu, W.; Wu, M.; Lu, C. Meticulous process monitoring with multiscale convolutional feature extraction. *J. Process Control* **2021**, *106*, 20–28. [CrossRef]
373. Li, S.; Luo, J.; Hu, Y. Nonlinear process modeling via unidimensional convolutional neural networks with self-attention on global and local inter-variable structures and its application to process monitoring. *ISA Trans.* **2022**, *121*, 105–118. [CrossRef] [PubMed]
374. Bao, Y.; Wang, B.; Guo, P.; Wang, J. Chemical process fault diagnosis based on a combined deep learning method. *Can. J. Chem. Eng.* **2022**, *100*, 54–66. [CrossRef]
375. Xing, J.; Xu, J. An Improved Convolutional Neural Network for Recognition of Incipient Faults. *IEEE Sens. J.* **2022**, *22*, 16314–16322. [CrossRef]

376. Zhang, S.; Wei, H.L.; Ding, J. An effective zero-shot learning approach for intelligent fault detection using 1D CNN. *Appl. Intell.* **2022**, *53*, 16041–16058. [CrossRef]
377. Guo, P.; Rao, S.; Hao, L.; Wang, J. Fault diagnosis of a semi-batch crystallization process through deep learning method. *Comput. Chem. Eng.* **2022**, *164*, 107807. [CrossRef]
378. Al-Wahaibi, S.S.S.; Abiola, S.; Chowdhury, M.A.; Lu, Q. Improving convolutional neural networks for fault diagnosis in chemical processes by incorporating global correlations. *Comput. Chem. Eng.* **2023**, *176*, 108289. [CrossRef]
379. Ma, F.; Ji, C.; Xu, M.; Wang, J.; Sun, W. Spatial Correlation Extraction for Chemical Process Fault Detection Using Image Enhancement Technique aided Convolutional Autoencoder. *Chem. Eng. Sci.* **2023**, *278*, 118900. [CrossRef]
380. Huang, Y.; Zhang, J.; Liu, R.; Zhao, S. Improving Accuracy and Interpretability of CNN-Based Fault Diagnosis through an Attention Mechanism. *Processes* **2023**, *11*, 3233. [CrossRef]
381. Liu, Y.; Jafarpour, B. Graph attention network with Granger causality map for fault detection and root cause diagnosis. *Comput. Chem. Eng.* **2024**, *180*, 108453. [CrossRef]
382. Souza, A.C.O.; Souza, M.B.d.; Silva, F.V.d. Enhancing fault detection and diagnosis systems for a chemical process: A study on convolutional neural networks and transfer learning. *Evol. Syst.* **2023**. [CrossRef]
383. Yu, J.; Zheng, X.; Wang, S. A deep autoencoder feature learning method for process pattern recognition. *J. Process Control* **2019**, *79*, 1–15. [CrossRef]
384. Hallgrimsson, A.D.; Niemann, H.H.; Lind, M. Unsupervised isolation of abnormal process variables using sparse autoencoders. *J. Process Control* **2021**, *99*, 107–119. [CrossRef]
385. Zhang, J.; Zhao, C. Condition-driven probabilistic adversarial autoencoder with nonlinear Gaussian feature learning for nonstationary process monitoring. *J. Process Control* **2022**, *117*, 140–156. [CrossRef]
386. Peng, C.; Kai, W.; Kun, Z.; Fanchao, M. Monitoring of wastewater treatment process based on multi-stage variational autoencoder. *Expert Syst. Appl.* **2022**, *207*, 117919. [CrossRef]
387. Zhu, J.; Shi, H.; Song, B.; Tao, Y.; Tan, S.; Zhang, T. Nonlinear process monitoring based on load weighted denoising autoencoder. *Measurement* **2021**, *171*, 108782. [CrossRef]
388. Zhang, C.; Yu, J.; Ye, L. Sparsity and manifold regularized convolutional auto-encoders-based feature learning for fault detection of multivariate processes. *Control Eng. Pract.* **2021**, *111*, 104811. [CrossRef]
389. Yin, J.; Yan, X. Stacked sparse autoencoders monitoring model based on fault-related variable selection. *Soft Comput.* **2021**, *25*, 3531–3543. [CrossRef]
390. Jang, K.; Pilario, K.E.S.; Lee, N.; Moon, I.; Na, J. Explainable Artificial Intelligence for Fault Diagnosis of Industrial Processes. *IEEE Trans. Ind. Inform.* **2023**, 1–8. [CrossRef]
391. Yang, J.; Wang, L. Nonlocal, local and global preserving stacked autoencoder based fault detection method for nonlinear process monitoring. *Chemom. Intell. Lab. Syst.* **2023**, *235*, 104758. [CrossRef]
392. Ruan, H.; Yu, J.; Shu, F.; Yang, X.; Li, Z. A Joint Stacked Autoencoder Approach with Silhouette Information for Industrial Fault Detection. *Processes* **2022**, *10*, 2408. [CrossRef]
393. Yu, F.; Liu, J.; Liu, D.; Wang, H.; Shang, L. Feature learning based on entropy estimation density peak clustering and stacked autoencoder for industrial process monitoring. *Can. J. Chem. Eng.* **2022**, *101*, 3998–4015. [CrossRef]
394. Liu, J.; Wu, J.; Xie, Y.; Jie, W.; Xu, P.; Tang, Z.; Yin, H. Toward robust process monitoring of complex process industries based on denoising sparse auto-encoder. *J. Ind. Inf. Integr.* **2022**, *30*, 100410. [CrossRef]
395. Cacciarelli, D.; Kulahci, M. A novel fault detection and diagnosis approach based on orthogonal autoencoders. *Comput. Chem. Eng.* **2022**, *163*, 107853. [CrossRef]
396. Liu, L.; Zhao, H.; Hu, Z. Graph dynamic autoencoder for fault detection. *Chem. Eng. Sci.* **2022**, *254*, 117637. [CrossRef]
397. Guo, Z.; Shang, C.; Ye, H. A novel similarity metric with application to big process data analytics. *Control Eng. Pract.* **2021**, *113*, 104843. [CrossRef]
398. Zhu, J.; Jiang, M.; Liu, Z. Fault Detection and Diagnosis in Industrial Processes with Variational Autoencoder: A Comprehensive Study. *Sensors* **2021**, *22*, 227. [CrossRef] [PubMed]
399. Gao, H.; Huang, W.; Gao, X.; Han, H. Decentralized adaptively weighted stacked autoencoder-based incipient fault detection for nonlinear industrial processes. *ISA Trans.* **2023**, *139*, 216–228. [CrossRef] [PubMed]
400. Zhang, Z.; Zhu, J.; Zhang, S.; Gao, F. Process monitoring using recurrent Kalman variational auto-encoder for general complex dynamic processes. *Eng. Appl. Artif. Intell.* **2023**, *123*, 106424. [CrossRef]
401. Li, D.; Liu, J.; Ma, X.; Jin, Q. Stacked supervised auto-encoder with graph regularization for feature extraction and fault classification in chemical processes. *J. Process Control* **2023**, *127*, 102999. [CrossRef]
402. Wang, C.; Tang, X.; Yu, J.; Yang, X.; Yan, X. Mechanistic block-based attention mechanism stacked autoencoder for describing typical unit connection industrial processes and their monitoring. *Can. J. Chem. Eng.* **2023**, *102*, 291–306. [CrossRef]
403. Aghaee, M.; Krau, S.; Tamer, M.; Budman, H. Unsupervised Fault Detection of Pharmaceutical Processes Using Long Short-Term Memory Autoencoders. *Ind. Eng. Chem. Res.* **2023**, *62*, 9773–9786. [CrossRef]
404. Pan, Z.; Wang, Y.; Cao, Y.; Gui, W. VAE-Based Interpretable Latent Variable Model for Process Monitoring. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–14. [CrossRef] [PubMed]
405. Zhu, X.; Yang, C.; Yang, C.; Gao, D.; Lou, S. An unsupervised fault monitoring framework for blast furnace: Information extraction enhanced GRU-GMM-autoencoder. *J. Process Control* **2023**, *130*, 103087. [CrossRef]

406. Du, W.; Yang, J.; Meng, G. Fault diagnosis for dynamic system based on the independent latent space reconstruction of generative adversarial network. *J. Process Control* **2023**, *125*, 28–40. [CrossRef]
407. Dzaferagic, M.; Marchetti, N.; Macaluso, I. Fault Detection and Classification in Industrial IoT in Case of Missing Sensor Data. *IEEE Internet Things J.* **2022**, *9*, 8892–8900. [CrossRef]
408. Ma, L.; Ding, Y.; Wang, Z.; Wang, C.; Ma, J.; Lu, C. An interpretable data augmentation scheme for machine fault diagnosis based on a sparsity-constrained generative adversarial network. *Expert Syst. Appl.* **2021**, *182*, 115234. [CrossRef]
409. Jang, K.; Kim, M.; Yoon, H.; Na, J.; Moon, I. Adversarial Autoencoder Based Nonlinear Process Monitoring. In *Computer Aided Chemical Engineering*; Türkay, M., Gani, R., Eds.; Elsevier: Amsterdam, The Netherlands, 2021; Volume 50, pp. 1195–1201. [CrossRef]
410. Lomov, I.; Lyubimov, M.; Makarov, I.; Zhukov, L.E. Fault detection in Tennessee Eastman process with temporal deep learning models. *J. Ind. Inf. Integr.* **2021**, *23*, 100216. [CrossRef]
411. Zhuo, Y.; Ge, Z. Gaussian Discriminative Analysis aided GAN for imbalanced big data augmentation and fault classification. *J. Process Control* **2020**, *92*, 271–287. [CrossRef]
412. Tian, W.; Liu, Z.; Li, L.; Zhang, S.; Li, C. Identification of abnormal conditions in high-dimensional chemical process based on feature selection and deep learning. *Chin. J. Chem. Eng.* **2020**, *28*, 1875–1883. [CrossRef]
413. Chai, Z.; Zhao, C. A Fine-Grained Adversarial Network Method for Cross-Domain Industrial Fault Diagnosis. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 1432–1442. [CrossRef]
414. Li, D.; Wang, H.; Zhou, J. Novelty Detection for Multimode Process Using GANs with Learning Disentangled Representation. In Proceedings of the 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 22–24 August 2020; pp. 2536–2541. [CrossRef]
415. Xu, Y.; Zou, X.; Ke, W.; Zhu, Q.; He, Y.; Zhang, M.; Zhang, Y. A novel imbalanced fault diagnosis method based on area identification conditional generative adversarial networks. *Can. J. Chem. Eng.* **2023**, *101*, 6944–6958. [CrossRef]
416. Deng, X.; Xiao, L.; Liu, X.; Zhang, X. One-dimensional residual GANomaly network-based deep feature extraction model for complex industrial system fault detection. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–13. [CrossRef]
417. Wei, Z.; Ji, X.; Zhou, L.; Dang, Y.; Dai, Y. A novel deep learning model based on target transformer for fault diagnosis of chemical process. *Process Saf. Environ. Prot.* **2022**, *167*, 480–492. [CrossRef]
418. Zhang, L.; Song, Z.; Zhang, Q.; Peng, Z. Generalized transformer in fault diagnosis of Tennessee Eastman process. *Neural Comput. Appl.* **2022**, *34*, 8575–8585. [CrossRef]
419. Yi, Y.; Zhao, H.; Hu, Z.; Peng, J. A local–global transformer for distributed monitoring of multi-unit nonlinear processes. *J. Process Control* **2023**, *122*, 13–26. [CrossRef]
420. Zhou, K.; Tong, Y.; Li, X.; Wei, X.; Huang, H.; Song, K.; Chen, X. Exploring global attention mechanism on fault detection and diagnosis for complex engineering processes. *Process Saf. Environ. Prot.* **2023**, *170*, 660–669. [CrossRef]
421. Bi, X.; Wu, D.; Xie, D.; Ye, H.; Zhao, J. Large-scale chemical process causal discovery from big data with transformer-based deep learning. *Process Saf. Environ. Prot.* **2023**, *173*, 163–177. [CrossRef]
422. Bai, Y.; Zhao, J. A novel transformer-based multi-variable multi-step prediction method for chemical process fault prognosis. *Process Saf. Environ. Prot.* **2023**, *169*, 937–947. [CrossRef]
423. Yi, S.; Zheng, S.; Yang, S.; Zhou, G.; He, J. Robust transformer-based anomaly detection for nuclear power data using maximum correntropy criterion. *Nucl. Eng. Technol.* 2023; *in press*. [CrossRef]
424. Guo, X.; Cui, X.; Cheng, C.; Lu, L. Transformer-based high-precision chemical process fault detection. In Proceedings of the 2023 2nd International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC), Mianyang, China, 11–13 August 2023; pp. 236–240. [CrossRef]
425. Zhu, Q.X.; Qian, Y.S.; Zhang, N.; He, Y.L.; Xu, Y. Multi-scale Transformer-CNN domain adaptation network for complex processes fault diagnosis. *J. Process Control* **2023**, *130*, 103069. [CrossRef]
426. Moody, J.M. Process Monitoring with Restricted Boltzmann Machines. Master's Thesis, Stellenbosch University, Stellenbosch, South Africa, 2014.
427. Zhang, Z.; Zhao, J. A deep belief network based fault diagnosis model for complex chemical processes. *Comput. Chem. Eng.* **2017**, *107*, 395–407. [CrossRef]
428. Wei, Y.; Weng, Z. Research on TE process fault diagnosis method based on DBN and Dropout. *Can. J. Chem. Eng.* **2020**, *98*, 1293–1306. [CrossRef]
429. Wang, Y.; Pan, Z.; Yuan, X.; Yang, C.; Gui, W. A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network. *ISA Trans.* **2020**, *96*, 457–467. [CrossRef] [PubMed]
430. Tang, P.; Peng, K.; Zhang, K.; Chen, Z.; Yang, X.; Li, L. A Deep Belief Network-based Fault Detection Method for Nonlinear Processes. *IFAC-PapersOnLine* **2018**, *51*, 9–14. [CrossRef]
431. Peng, P.; Wu, Y.; Zhang, Y.; Zhang, H. A Novel Fault Detection and Diagnosis Method Based on Gaussian-Bernoulli Restricted Boltzmann Machine. In Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 6–9 October 2019; pp. 1349–1354. [CrossRef]
432. Yang, J.; Bao, W.; Liu, Y.; Li, X.; Wang, J.; Niu, Y. A pairwise graph regularized constraint based on deep belief network for fault diagnosis. *Digit. Signal Process.* **2021**, *108*, 102868. [CrossRef]

433. Ying, Y.; Li, Z.; Yang, M.; Du, W. Multimode Operating Performance Visualization and Nonoptimal Cause Identification. *Processes* **2020**, *8*, 123. [CrossRef]
434. Yan, S.; Huang, J.; Yan, X. Monitoring of quality-relevant and quality-irrelevant blocks with characteristic-similar variables based on self-organizing map and kernel approaches. *J. Process Control* **2019**, *73*, 103–112. [CrossRef]
435. Lu, W.; Yan, X. Deep fisher autoencoder combined with self-organizing map for visual industrial process monitoring. *J. Manuf. Syst.* **2020**, *56*, 241–251. [CrossRef]
436. Chen, X.; Yan, X. Using improved self-organizing map for fault diagnosis in chemical industry process. *Chem. Eng. Res. Des.* **2012**, *90*, 2262–2277. [CrossRef]
437. Ng, Y.S.; Srinivasan, R. Multivariate Temporal Data Analysis Using Self-Organizing Maps. 2. Monitoring and Diagnosis of Multistate Operations. *Ind. Eng. Chem. Res.* **2008**, *47*, 7758–7771. [CrossRef]
438. Yu, H.; Khan, F.; Garaniya, V.; Ahmad, A. Self-Organizing Map Based Fault Diagnosis Technique for Non-Gaussian Processes. *Ind. Eng. Chem. Res.* **2014**, *53*, 8831–8843. [CrossRef]
439. von Birgelen, A.; Buratti, D.; Mager, J.; Niggemann, O. Self-Organizing Maps for Anomaly Localization and Predictive Maintenance in Cyber-Physical Production Systems. *Procedia CIRP* **2018**, *72*, 480–485. [CrossRef]
440. Lu, W.; Yan, X. Deep model based on mode elimination and Fisher criterion combined with self-organizing map for visual multimodal chemical process monitoring. *Inf. Sci.* **2021**, *562*, 13–27. [CrossRef]
441. Silva, R.G.; Wilcox, S.J. Feature evaluation and selection for condition monitoring using a self-organizing map and spatial statistics. *Artif. Intell. Eng. Des. Anal. Manuf.* **2019**, *33*, 1–10. [CrossRef]
442. Ayoubi, M.; Isermann, R. Neuro-fuzzy systems for diagnosis. *Fuzzy Sets Syst.* **1997**, *89*, 289–307. [CrossRef]
443. Evsukoff, A.; Gentil, S. Recurrent neuro-fuzzy system for fault detection and isolation in nuclear reactors. *Adv. Eng. Inform.* **2005**, *19*, 55–66. [CrossRef]
444. Khalid, H.M.; Khoukhi, A.; Al-Sunni, F.M. Fault detection and classification using Kalman filter and genetic neuro-fuzzy systems. In Proceedings of the 2011 Annual Meeting of the North American Fuzzy Information Processing Society, El Paso, TX, USA, 18–20 March 2011. [CrossRef]
445. Lau, C.K.; Ghosh, K.; Hussain, M.A.; Che Hassan, C.R. Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS. *Chemom. Intell. Lab. Syst.* **2013**, *120*, 1–14. [CrossRef]
446. Gholizadeh, M.; Yazdizadeh, A.; Mohammad-Bagherpour, H. Fault detection and identification using combination of EKF and neuro-fuzzy network applied to a chemical process (CSTR). *Pattern Anal. Appl.* **2019**, *22*, 359–373. [CrossRef]
447. Shabanian, M.; Montazeri, M. A neuro-fuzzy online fault detection and diagnosis algorithm for nonlinear and dynamic systems. *Int. J. Control. Autom. Syst.* **2011**, *9*, 665–670. [CrossRef]
448. Subbaraj, P.; Kannapiran, B. Fault detection and diagnosis of pneumatic valve using Adaptive Neuro-Fuzzy Inference System approach. *Appl. Soft Comput.* **2014**, *19*, 362–371. [CrossRef]
449. Salahshoor, K.; Kordestani, M.; Khoshro, M.S. Fault detection and diagnosis of an industrial steam turbine using fusion of SVM (support vector machine) and ANFIS (adaptive neuro-fuzzy inference system) classifiers. *Energy* **2010**, *35*, 5472–5482. [CrossRef]
450. Korbicz, J.; Kowal, M. Neuro-fuzzy networks and their application to fault detection of dynamical systems. *Eng. Appl. Artif. Intell.* **2007**, *20*, 609–617. [CrossRef]
451. Palade, V.; Patton, R.J.; Uppal, F.J.; Quevedo, J.; Daley, S. Fault diagnosis of an industrial gas turbine using neuro-fuzzy methods. *IFAC Proc. Vol.* **2002**, *35*, 471–476. [CrossRef]
452. Boldt, F.d.A.; Rauber, T.W.; Varejão, F.M. Cascade Feature Selection and ELM for automatic fault diagnosis of the Tennessee Eastman process. *Neurocomputing* **2017**, *239*, 238–248. [CrossRef]
453. Zou, W.; Xia, Y.; Li, H. Fault Diagnosis of Tennessee-Eastman Process Using Orthogonal Incremental Extreme Learning Machine Based on Driving Amount. *IEEE Trans. Cybern.* **2018**, *48*, 3403–3410. [CrossRef] [PubMed]
454. Zhang, H.; Deng, X.; Zhang, Y.; Hou, C.; Li, C.; Xin, Z. Nonlinear Process Monitoring Based on Global Preserving Unsupervised Kernel Extreme Learning Machine. *IEEE Access* **2019**, *7*, 106053–106064. [CrossRef]
455. Yongyong, H.; Xiaoqiang, Z. Sparse representation preserving embedding based on extreme learning machine for process monitoring. *Trans. Inst. Meas. Control* **2020**, *42*, 014233121989893. [CrossRef]
456. Peng, X.; Li, H.; Yuan, F.; Razul, S.G.; Chen, Z.; Lin, Z. An extreme learning machine for unsupervised online anomaly detection in multivariate time series. *Neurocomputing* **2022**, *501*, 596–608. [CrossRef]
457. Cavalcanti, F.M.; Kozonoe, C.E.; Pacheco, K.A.; Alves, R.M.d.B. Application of Artificial Neural Networks to Chemical and Process Engineering. In *Artificial Neural Networks and Deep Learning—Applications and Perspective*; IntechOpen: Rijeka, Croatia, 2021. [CrossRef]
458. Jiang, S.; Zavala, V.M. Convolutional neural nets in chemical engineering: Foundations, computations, and applications. *AIChE J.* **2021**, *67*, e17282. [CrossRef]
459. Sun, L.; Liang, F.; Cui, W. Artificial Neural Network and Its Application Research Progress in Chemical Process. *arXiv* **2021**, arXiv:2110.09021.
460. He, Q.P.; Wang, J. Fault Detection Using the k-Nearest Neighbor Rule for Semiconductor Manufacturing Processes. *IEEE Trans. Semicond. Manuf.* **2007**, *20*, 345–354. [CrossRef]
461. Stockmann, M.; Haber, R.; Schmitz, U. Source identification of plant-wide faults based on k nearest neighbor time delay estimation. *J. Process Control* **2012**, *22*, 583–598. [CrossRef]

462. Wang, G.; Liu, J.; Li, Y. Fault diagnosis using kNN reconstruction on MRI variables: Fault diagnosis. *J. Chemom.* **2015**, *29*, 399–410. [CrossRef]
463. Zhou, Z.; Wen, C.; Yang, C. Fault Isolation Based On k-Nearest Neighbor Rule For Industrial Processes. *IEEE Trans. Ind. Electron.* **2016**, *63*, 2578–2586. [CrossRef]
464. Senin, P. *Dynamic Time Warping Algorithm Review*; University of Hawaii at Manoa: Honolulu, HI, USA, 2008.
465. Berndt, D.J.; Clifford, J. Using Dynamic Time Warping to Find Patterns in Time Series. In Proceedings of the AAAIWS'94: Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 31 July–1 August 1994; pp. 359–370.
466. Muller, M. Dynamic Time Warping. In *Information Retrieval for Music and Motion*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 69–84. [CrossRef]
467. Si, Y.; Chen, Z.; Sun, J.; Zhang, D.; Qian, P. A Data-Driven Fault Detection Framework Using Mahalanobis Distance Based Dynamic Time Warping. *IEEE Access* **2020**, *8*, 108359–108370. [CrossRef]
468. Feng, J.; Li, K. MRS kNN fault detection method for multirate sampling process based variable grouping threshold. *J. Process Control* **2020**, *85*, 149–158. [CrossRef]
469. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the COLT '92: Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; pp. 144–152. [CrossRef]
470. Tax, D.M.; Duin, R.P. Support vector domain description. *Pattern Recognit. Lett.* **1999**, *20*, 1191–1199. [CrossRef]
471. Scholkopf, B.; Platt, J.C.; Shawe-Taylor, J.; Smola, A.J.; Williamson, R.C. Estimating the Support of a High-Dimensional Distribution. *Neural Comput.* **2001**, *13*, 1443–1471. [CrossRef]
472. Xiao, Y.; Wang, H.; Zhang, L.; Xu, W. Two methods of selecting Gaussian kernel parameters for one-class SVM and their application to fault detection. *Knowl.-Based Syst.* **2014**, *59*, 75–84. [CrossRef]
473. Mahadevan, S.; Shah, S.L. Fault detection and diagnosis in process data using one-class support vector machines. *J. Process Control* **2009**, *19*, 1627–1639. [CrossRef]
474. Yang, C.; Hou, J. Fed-batch fermentation penicillin process fault diagnosis and detection based on support vector machine. *Neurocomputing* **2016**, *190*, 117–123. [CrossRef]
475. Lee, S.; Kim, S.B. Time-adaptive support vector data description for nonstationary process monitoring. *Eng. Appl. Artif. Intell.* **2018**, *68*, 18–31. [CrossRef]
476. Yuan, P.; Mao, Z.; Wang, B. A pruned support vector data description -based outlier detection method: Applied to robust process monitoring. *Trans. Inst. Meas. Control* **2020**, *42*, 014233122090595. [CrossRef]
477. Zhang, Y.; Li, X. Two step support vector data description for dynamic, non linear, and non Gaussian processes monitoring. *Can. J. Chem. Eng.* **2020**, *98*, 2109–2124. [CrossRef]
478. Zhang, C.; Dong, J.; Peng, K.; Zhang, X. A novel quality-related process monitoring method for multi-unit industrial processes under incomplete data conditions. *Can. J. Chem. Eng.* **2022**, *101*, 1485–1498. [CrossRef]
479. Yin, Z.; Hou, J. Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing* **2016**, *174*, 643–650. [CrossRef]
480. Widodo, A.; Yang, B.S. Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process.* **2007**, *21*, 2560–2574. [CrossRef]
481. Grus, J. *Data Science from Scratch: First Principles with Python*, 1st ed.; O'Reilly: Sebastopol, CA, USA, 2015.
482. Guh, R.S.; Shiue, Y.R. An effective application of decision tree learning for on-line detection of mean shifts in multivariate control charts. *Comput. Ind. Eng.* **2008**, *55*, 475–493. [CrossRef]
483. He, S.G.; He, Z.; Wang, G.A. Online monitoring and fault identification of mean shifts in bivariate processes using decision tree learning techniques. *J. Intell. Manuf.* **2013**, *24*, 25–34. [CrossRef]
484. Aldrich, C.; Auret, L. Fault detection and diagnosis with random forest feature extraction and variable importance methods. *IFAC Proc. Vol.* **2010**, *43*, 79–86. [CrossRef]
485. Auret, L.; Aldrich, C. Unsupervised Process Fault Detection with Random Forests. *Ind. Eng. Chem. Res.* **2010**, *49*, 9184–9194. [CrossRef]
486. Puggini, L.; Doyle, J.; McLoone, S. Fault Detection using Random Forest Similarity Distance. *IFAC-PapersOnLine* **2015**, *48*, 583–588. [CrossRef]
487. Liu, Y.; Ge, Z. Weighted random forests for fault classification in industrial processes with hierarchical clustering model selection. *J. Process Control* **2018**, *64*, 62–70. [CrossRef]
488. Fan, S.K.S.; Hsu, C.Y.; Tsai, D.M.; He, F.; Cheng, C.C. Data-Driven Approach for Fault Detection and Diagnostic in Semiconductor Manufacturing. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 1925–1936. [CrossRef]
489. Gajjar, S.; Kulahci, M.; Palazoglu, A. Least Squares Sparse Principal Component Analysis and Parallel Coordinates for Real-Time Process Monitoring. *Ind. Eng. Chem. Res.* **2020**, *59*, 15656–15670. [CrossRef]
490. Liu, Y.; Ge, Z. Deep ensemble forests for industrial fault classification. *IFAC J. Syst. Control* **2019**, *10*, 100071. [CrossRef]
491. Xu, Y.; Cong, K.; Zhu, Q.; He, Y. A novel AdaBoost ensemble model based on the reconstruction of local tangent space alignment and its application to multiple faults recognition. *J. Process Control* **2021**, *104*, 158–167. [CrossRef]

492. Shrivastava, R. Comparative study of boosting and bagging based methods for fault detection in a chemical process. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 25–27 March 2021; pp. 674–679. [CrossRef]
493. Nogueira, J.d.N.P.; Melo, P.A.; de Souza, M.B., Jr. Faulty Scenarios in Sour Water Treatment Units: Simulation and AI-Based Diagnosis. *Process Saf. Environ. Prot.* **2022**, *165*, 716–727. [CrossRef]
494. Zheng, J.; Liu, Y.; Ge, Z. Dynamic ensemble selection based improved random forests for fault classification in industrial processes. *IFAC J. Syst. Control* **2022**, *20*, 100189. [CrossRef]
495. Sahraoui, M.A.; Rahmoune, C.; Zair, M.; Gougam, F.; Damou, A. Enhancing fault diagnosis of undesirable events in oil & gas systems: A machine learning approach with new criteria for stability analysis and classification accuracy. *Proc. Inst. Mech. Eng. Part J. Process Mech. Eng.* **2023**, 09544089231213778. [CrossRef]
496. Ayesha, S.; Hanif, M.K.; Talib, R. Overview and comparative study of dimensionality reduction techniques for high dimensional data. *Inf. Fusion* **2020**, *59*, 44–58. [CrossRef]
497. Gisbrecht, A.; Hammer, B. Data visualization by nonlinear dimensionality reduction. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2015**, *5*, 51–73. [CrossRef]
498. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
499. McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* **2020**, arXiv:1802.03426. [CrossRef]
500. He, F.; Wang, C.; Fan, S.K.S. Nonlinear fault detection of batch processes based on functional kernel locality preserving projections. *Chemom. Intell. Lab. Syst.* **2018**, *183*, 79–89. [CrossRef]
501. López, C.A.M.; Bhonsale, S.; Peeters, K.; Van Impe, J.F.M. Manifold Learning and Clustering for Automated Phase Identification and Alignment in Data Driven Modeling of Batch Processes. *Front. Chem. Eng.* **2020**, *2*, 582126. [CrossRef]
502. Reis, M.S.; Rendall, R.; Rato, T.J.; Martins, C.; Delgado, P. Improving the sensitivity of statistical process monitoring of manifolds embedded in high-dimensional spaces: The truncated-Q statistic. *Chemom. Intell. Lab. Syst.* **2021**, *215*, 104369. [CrossRef]
503. Zhang, M.; Ge, Z.; Song, Z.; Fu, R. Global–Local Structure Analysis Model and Its Application for Fault Detection and Identification. *Ind. Eng. Chem. Res.* **2011**, *50*, 6837–6848. [CrossRef]
504. Yu, J. Process monitoring through manifold regularization-based GMM with global/local information. *J. Process Control* **2016**, *45*, 84–99. [CrossRef]
505. Luo, L.; Bao, S.; Mao, J.; Tang, D. Nonlocal and local structure preserving projection and its application to fault detection. *Chemom. Intell. Lab. Syst.* **2016**, *157*, 177–188. [CrossRef]
506. Zhan, C.; Li, S.; Yang, Y. Improved process monitoring based on global–local manifold analysis and statistical local approach for industrial process. *J. Process Control* **2019**, *75*, 107–119. [CrossRef]
507. Fu, Y.; Luo, C. Joint Structure Preserving Embedding Model and Its Application for Process Monitoring. *Ind. Eng. Chem. Res.* **2019**, *58*, 20667–20679. [CrossRef]
508. Cui, P.; Wang, X.; Yang, Y. Nonparametric manifold learning approach for improved process monitoring. *Can. J. Chem. Eng.* **2022**, *100*, 67–89. [CrossRef]
509. Li, Y.; Ma, F.; Ji, C.; Wang, J.; Sun, W. Fault Detection Method Based on Global-Local Marginal Discriminant Preserving Projection for Chemical Process. *Processes* **2022**, *10*, 122. [CrossRef]
510. Olshausen, B.A.; Field, D.J. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vis. Res.* **1997**, *37*, 3311–3325. [CrossRef] [PubMed]
511. Mairal, J.; Bach, F.; Ponce, J.; Sapiro, G.; Zisserman, A. Supervised Dictionary Learning. *arXiv* **2008**, arXiv:0809.3083.
512. Wu, L.; Chen, X.; Peng, Y.; Ye, Q.; Jiao, J. Fault detection and diagnosis based on sparse representation classification (SRC). In Proceedings of the 2012 IEEE International Conference on Robotics and Biomimetics (ROBIO), Guangzhou, China, 11–14 December 2012; pp. 926–931. [CrossRef]
513. Ren, L.; Lv, W. Fault Detection via Sparse Representation for Semiconductor Manufacturing Processes. *IEEE Trans. Semicond. Manuf.* **2014**, *27*, 252–259. [CrossRef]
514. Ning, C.; Chen, M.; Zhou, D. Sparse Contribution Plot for Fault Diagnosis of Multimodal Chemical Processes. *IFAC-PapersOnLine* **2015**, *48*, 619–626. [CrossRef]
515. Peng, X.; Tang, Y.; Du, W.; Qian, F. Multimode Process Monitoring and Fault Detection: A Sparse Modeling and Dictionary Learning Method. *IEEE Trans. Ind. Electron.* **2017**, *64*, 4866–4875. [CrossRef]
516. Huang, K.; Wen, H.; Ji, H.; Cen, L.; Chen, X.; Yang, C. Nonlinear process monitoring using kernel dictionary learning with application to aluminum electrolysis process. *Control Eng. Pract.* **2019**, *89*, 94–102. [CrossRef]
517. Yang, C.; Zhou, L.; Huang, K.; Ji, H.; Long, C.; Chen, X.; Xie, Y. Multimode process monitoring based on robust dictionary learning with application to aluminium electrolysis process. *Neurocomputing* **2019**, *332*, 305–319. [CrossRef]
518. Huang, K.; Wu, Y.; Wen, H.; Liu, Y.; Yang, C.; Gui, W. Distributed dictionary learning for high-dimensional process monitoring. *Control Eng. Pract.* **2020**, *98*, 104386. [CrossRef]
519. Huang, K.; Wei, K.; Li, Y.; Yang, C. Distributed dictionary learning for industrial process monitoring with big data. *Appl. Intell.* **2021**. [CrossRef]
520. Huang, K.; Tao, Z.; Liu, Y.; Sun, B.; Yang, C.; Gui, W.; Hu, S. Adaptive Multimode Process Monitoring Based on Mode-Matching and Similarity-Preserving Dictionary Learning. *IEEE Trans. Cybern.* **2022**, *53*, 3974–3987. [CrossRef] [PubMed]

521. Yang, C.; Zhang, J.; Wu, D.; Huang, K.; Gui, W. Variable partition based parallel dictionary learning for linearity and nonlinearity coexisting dynamic process monitoring. *Control Eng. Pract.* **2024**, *142*, 105750. [CrossRef]
522. Jun, H.B.; Kim, D. A Bayesian network-based approach for fault analysis. *Expert Syst. Appl.* **2017**, *81*, 332–348. [CrossRef]
523. Cai, Z.; Sun, S.; Si, S.; Wang, N. Research of failure prediction Bayesian network model. In Proceedings of the 2009 16th International Conference on Industrial Engineering and Engineering Management, Beijing, China, 21–23 October 2009; pp. 2021–2025. [CrossRef]
524. Lampis, M.; Andrews, J.D. Bayesian belief networks for system fault diagnostics. *Qual. Reliab. Eng. Int.* **2009**, *25*, 409–426. [CrossRef]
525. Liu, H.; Buvat, J.C.; Estel, L.; Polaert, I. Bayesian Network Method for Fault Diagnosis in a Continuous Tubular Reactor. *Chem. Prod. Process Model.* **2010**, *5*, 28. [CrossRef]
526. Verron, S.; Li, J.; Tiplica, T. Fault detection and isolation of faults in a multivariate process with Bayesian network. *J. Process Control* **2010**, *20*, 902–911. [CrossRef]
527. Yu, J.; Rashid, M.M. A novel dynamic bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis. *AIChE J.* **2013**, *59*, 2348–2365. [CrossRef]
528. Liu, J.; Wong, D.S.H.; Chen, D.S. Bayesian filtering of the smearing effect: Fault isolation in chemical process monitoring. *J. Process Control* **2014**, *24*, 1–21. [CrossRef]
529. Chen, G.; Ge, Z. Robust Bayesian networks for low-quality data modeling and process monitoring applications. *Control Eng. Pract.* **2020**, *97*, 104344. [CrossRef]
530. Lucke, M.; Stief, A.; Chioua, M.; Ottewill, J.R.; Thornhill, N.F. Fault detection and identification combining process measurements and statistical alarms. *Control Eng. Pract.* **2020**, *94*, 104195. [CrossRef]
531. Nhat, D.M.; Venkatesan, R.; Khan, F. Data-driven Bayesian network model for early kick detection in industrial drilling process. *Process Saf. Environ. Prot.* **2020**, *138*, 130–138. [CrossRef]
532. Liu, N.; Hu, M.; Wang, J.; Ren, Y.; Tian, W. Fault detection and diagnosis using Bayesian network model combining mechanism correlation analysis and process data: Application to unmonitored root cause variables type faults. *Process Saf. Environ. Prot.* **2022**, *164*, 15–29. [CrossRef]
533. Yang, W.T.; Reis, M.S.; Borodin, V.; Juge, M.; Roussy, A. An interpretable unsupervised Bayesian network model for fault detection and diagnosis. *Control Eng. Pract.* **2022**, *127*, 105304. [CrossRef]
534. Ge, Z.; Song, Z. Multimode process monitoring based on Bayesian method. *J. Chemom.* **2009**, *23*, 636–650. [CrossRef]
535. Wang, Y.; Jiang, Q.; Fu, J. Efficient Monitoring of Nonlinear Chemical Processes based on Fault-Relevant Kernel Principal Component Subspace Construction and Bayesian Inference. *J. Chem. Eng. Jpn.* **2017**, *50*, 648–656. [CrossRef]
536. Jiang, Q.; Yan, X.; Li, J. PCA-ICA Integrated with Bayesian Method for Non-Gaussian Fault Diagnosis. *Ind. Eng. Chem. Res.* **2016**, *55*, 4979–4986. [CrossRef]
537. Raveendran, R.; Huang, B. Conjugate exponential family graphical models in process monitoring: A tutorial review. *Chemom. Intell. Lab. Syst.* **2020**, *204*, 104095. [CrossRef]
538. Lemos, A.; Caminhas, W.; Gomide, F. Adaptive fault detection and diagnosis using an evolving fuzzy classifier. *Inf. Sci.* **2013**, *220*, 64–85. [CrossRef]
539. Ramos, A.R.; Corona, C.C.; Verdegay, J.L.; da Silva Neto, A.J.; Llanes-Santiago, O. An approach for fault diagnosis using a novel hybrid fuzzy clustering algorithm. In Proceedings of the 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Rio de Janeiro, Brazil, 8–13 July 2018. [CrossRef]
540. Ribeiro, K.; Fontes, C.H.; de Melo, G.J.A. Genetic algorithm-based fuzzy clustering applied to multivariate time series. *Evol. Intell.* **2021**, *14*, 1547–1563. [CrossRef]
541. Elshenawy, L.M.; Chakour, C.; Mahmoud, T.A. Fault detection and diagnosis strategy based on k-nearest neighbors and fuzzy C-means clustering algorithm for industrial processes. *J. Frankl. Inst.* **2022**, *359*, 7115–7139. [CrossRef]
542. Ezugwu, A.E.; Ikotun, A.M.; Oyelade, O.O.; Abualigah, L.; Agushaka, J.O.; Eke, C.I.; Akinyelu, A.A. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Eng. Appl. Artif. Intell.* **2022**, *110*, 104743. [CrossRef]
543. Barragan, J.F.; Fontes, C.H.; Embiruçu, M. A wavelet-based clustering of multivariate time series using a Multiscale SPCA approach. *Comput. Ind. Eng.* **2016**, *95*, 144–155. [CrossRef]
544. Fontes, C.H.; Budman, H. A hybrid clustering approach for multivariate time series—A case study applied to failure analysis in a gas turbine. *ISA Trans.* **2017**, *71*, 513–529. [CrossRef]
545. Hammer, P.L. Partially defined Boolean functions and cause-effect relationships. In Proceedings of the the International Conference on Multi-Attribute Decision Making via OR-Based Expert Systems, University of Passau, Passau, Germany, 20–27 April 1986.
546. Lejeune, M.; Lozin, V.; Lozina, I.; Ragab, A.; Yacout, S. Recent advances in the theory and practice of Logical Analysis of Data. *Eur. J. Oper. Res.* **2019**, *275*, 1–15. [CrossRef]
547. Mortada, M.A.; Carroll, T.; Yacout, S.; Lakis, A. Rogue components: Their effect and control using logical analysis of data. *J. Intell. Manuf.* **2012**, *23*, 289–302. [CrossRef]

548. Bai, X.; Tan, J.; Wang, X. Fault Diagnosis and Knowledge Extraction Using Fast Logical Analysis of Data with Multiple Rules Discovery Ability. In Proceedings of the 2nd International Conference on Intelligence Science (ICIS), Beijing, China, 2–5 November 2018; IFIP Advances in Information and Communication Technology; Shi, Z., Pennartz, C., Huang, T., Eds.; Springer: Cham, Switzerland, 2018; pp. 412–421. [CrossRef]
549. Ragab, A.; El Koujok, M.; Ghezzaz, H.; Amazouz, M.; Ouali, M.S.; Yacout, S. Deep understanding in industrial processes by complementing human expertise with interpretable patterns of machine learning. *Expert Syst. Appl.* **2019**, *122*, 388–405. [CrossRef]
550. Ragab, A.; El-Koujok, M.; Poulin, B.; Amazouz, M.; Yacout, S. Fault diagnosis in industrial chemical processes using interpretable patterns based on Logical Analysis of Data. *Expert Syst. Appl.* **2018**, *95*, 368–383. [CrossRef]
551. Ragab, A.; El-koujok, M.; Amazouz, M.; Yacout, S. Fault detection and diagnosis in the Tennessee Eastman Process using interpretable knowledge discovery. In Proceedings of the 2017 Annual Reliability and Maintainability Symposium (RAMS), Orlando, FL, USA, 23–26 January 2017. [CrossRef]
552. Angelov, P.P. *Autonomous Learning Systems: From Data Streams to Knowledge in Real-Time*; Wiley John Wiley & Sons, Ltd.: Chichester, West Sussex, UK, 2013.
553. Angelov, P. Anomaly detection based on eccentricity analysis. In Proceedings of the 2014 IEEE Symposium on Evolving and Autonomous Learning Systems (EALS), Orlando, FL, USA, 9–12 December 2014; pp. 1–8. [CrossRef]
554. Angelov, P.; Gu, X.; Kangin, D.; Principe, J. Empirical data analysis: A new tool for data analytics. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; pp. 000052–000059. [CrossRef]
555. Angelov, P.; Gu, X.; Kangin, D. Empirical Data Analytics. *Int. J. Intell. Syst.* **2017**, *32*, 1261–1284. [CrossRef]
556. Angelov, P.P.; Gu, X. *Empirical Approach to Machine Learning*; *Studies in Computational Intelligence*; Springer International Publishing: Cham, Switzerland, 2019; Volume 800. [CrossRef]
557. Costa, B.S.J.; Angelov, P.P.; Guedes, L.A. Real-Time Fault Detection Using Recursive Density Estimation. *J. Control. Autom. Electr. Syst.* **2014**, *25*, 428–437. [CrossRef]
558. Costa, B.S.J.; Angelov, P.P.; Guedes, L.A. Fully unsupervised fault detection and identification based on recursive density estimation and self-evolving cloud-based classifier. *Neurocomputing* **2015**, *150*, 289–303. [CrossRef]
559. Costa, B.S.J.; Bezerra, C.G.; Guedes, L.A.; Angelov, P.P. Online fault detection based on Typicality and Eccentricity Data Analytics. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–17 July 2015. [CrossRef]
560. Bezerra, C.G.; Costa, B.S.J.; Guedes, L.A.; Angelov, P.P. An evolving approach to unsupervised and Real-Time fault detection in industrial processes. *Expert Syst. Appl.* **2016**, *63*, 134–144. [CrossRef]
561. Germano, A.L.; Guedes, L.A.; Costa, B.S.J.; Bezerra, G. Detecção de falhas no processo Tennessee Eastman Utilizando Métricas de Tipicidade e Excentricidade. In Proceedings of the XXI Congresso Brasileiro de Automática—CBA2016, Espírito Santo, Brasil, 3–7 October 2016.
562. Ziaei-Halimejani, H.; Kouzbour, S.; Zarghami, R.; Sotudeh-Gharebagh, R.; Gourich, B.; Mostoufi, N.; Stiriba, Y. Monitoring of the bubble columns hydrodynamics by recurrence quantification data analysis. *Chem. Eng. Res. Des.* **2021**, *171*, 100–110. [CrossRef]
563. Ziaei-Halimejani, H.; Nazemzadeh, N.; Zarghami, R.; Gernaey, K.V.; Andersson, M.P.; Mansouri, S.S.; Mostoufi, N. Fault Diagnosis of Chemical Processes based on Joint Recurrence Quantification Analysis. *Comput. Chem. Eng.* **2021**, *155*, 107549. [CrossRef]
564. Ziaei-Halimejani, H.; Zarghami, R.; Mansouri, S.S.; Mostoufi, N. Data-Driven Fault Diagnosis of Chemical Processes Based on Recurrence Plots. *Ind. Eng. Chem. Res.* **2021**, *60*, 3038–3055. [CrossRef]
565. Ziaei-Halimejani, H.; Zarghami, R.; Mostoufi, N. Joint recurrence based root cause analysis of nonlinear multivariate chemical processes. *J. Process Control* **2021**, *103*, 19–33. [CrossRef]
566. Bardinas, J. Characterisation of Dynamic Process Systems by Use of Recurrence Texture Analysis. Ph.D. Thesis, Curtin University, Bentley, Australia, 2018.
567. Bardinas, J.; Aldrich, C.; Napier, L. Predicting the Operating States of Grinding Circuits by Use of Recurrence Texture Analysis of Time Series Data. *Processes* **2018**, *6*, 17. [CrossRef]
568. Kok, T.L.; Aldrich, C.; Zabiri, H.; Taqvi, S.A.A.; Olivier, J. Application of unthresholded recurrence plots and texture analysis for industrial loops with faulty valves. *Soft Comput.* **2022**, *26*, 10477–10492. [CrossRef]
569. Melo, A.; Fernando, F.; Câmara, M.M.; Pinto, J.C. Distance matrix patterns for visual and interpretable process data analytics. *Ind. Eng. Chem. Res.* **2023**, *62*, 13889–13901. [CrossRef]
570. Pearson, R. Exploring process data. *J. Process Control* **2001**, *11*, 179–194. [CrossRef]
571. Abonyi, J. Application of Exploratory Data Analysis to Historical Process Data of Polyethylene Production. *Bulg. J. Ind. Chem.* **2007**, *35*, 85–93.
572. Xu, S.; Lu, B.; Baldea, M.; Edgar, T.F.; Wojsznis, W.; Blevins, T.; Nixon, M. Data cleaning in the process industries. *Rev. Chem. Eng.* **2015**, *31*, 453–490. [CrossRef]
573. Wang, J.; He, Q.P. Multivariate Statistical Process Monitoring Based on Statistics Pattern Analysis. *Ind. Eng. Chem. Res.* **2010**, *49*, 7858–7869. [CrossRef]

574. He, Q.P.; Wang, J. Statistics pattern analysis: A new process monitoring framework and its application to semiconductor batch processes. *AIChE J.* **2011**, *57*, 107–121. [CrossRef]
575. Lucke, M.; Mei, X.; Stief, A.; Chioua, M.; Thornhill, N.F. Variable Selection for Fault Detection and Identification based on Mutual Information of Alarm Series. *IFAC-PapersOnLine* **2019**, *52*, 673–678. [CrossRef]
576. Rauber, T.; Boldt, F.; Munaro, C. Feature Selection for Multivariate Contribution Analysis in Fault Detection and Isolation. *J. Frankl. Inst.* **2020**, *357*, 6294–6320. [CrossRef]
577. Kwak, S.; Ma, Y.; Huang, B. Extracting nonstationary features for process data analytics and application in fouling detection. *Comput. Chem. Eng.* **2020**, *135*, 106762. [CrossRef]
578. Wu, D.; Zhou, D.; Zhang, J.; Chen, M. Multimode process monitoring based on fault dependent variable selection and moving window-negative log likelihood probability. *Comput. Chem. Eng.* **2020**, *136*, 106787. [CrossRef]
579. Shah, D.; Wang, J.; He, Q.P. Feature engineering in big data analytics for IoT-enabled smart manufacturing – Comparison between deep learning and statistical learning. *Comput. Chem. Eng.* **2020**, *141*, 106970. [CrossRef]
580. Clavijo, N.; Melo, A.; Soares, R.M.; Campos, L.F.d.O.; Lemos, T.; Câmara, M.M.; Anzai, T.K.; Diehl, F.C.; Thompson, P.H.; Pinto, J.C. Variable Selection for Fault Detection Based on Causal Discovery Methods: Analysis of an Actual Industrial Case. *Processes* **2021**, *9*, 544. [CrossRef]
581. Siang, L.C.; Elnawawi, S.; Rippon, L.D.; O'Connor, D.L.; Gopaluni, R.B. Data Quality Over Quantity: Pitfalls and Guidelines for Process Analytics. *arXiv* **2023**, arXiv:2211.06440.
582. Thibault, E.; Chioua, M.; McKay, M.; Korbel, M.; Patience, G.S.; Stuart, P.R. Experimental methods in chemical engineering: Data processing and data usage in decision-making. *Can. J. Chem. Eng.* **2023**, *101*, 6055–6078. [CrossRef]
583. Li, W.; Zhao, C.; Gao, F. Linearity Evaluation and Variable Subset Partition Based Hierarchical Process Modeling and Monitoring. *IEEE Trans. Ind. Electron.* **2018**, *65*, 2683–2692. [CrossRef]
584. Thomas, M.C.; Zhu, W.; Romagnoli, J.A. Data mining and clustering in chemical process databases for monitoring and knowledge discovery. *J. Process Control* **2018**, *67*, 160–175. [CrossRef]
585. Pereira Parente, A.; de Souza Jr., M.B.; Valdman, A.; Mattos Folly, R.O. Data Augmentation Applied to Machine Learning-Based Monitoring of a Pulp and Paper Process. *Processes* **2019**, *7*, 958. [CrossRef]
586. Offermans, T.; Szymanska, E.; Buydens, L.M.; Jansen, J.J. Synchronizing process variables in time for industrial process monitoring and control. *Comput. Chem. Eng.* **2020**, *140*, 106938. [CrossRef]
587. Rhyu, J.; Bozinovski, D.; Dubs, A.B.; Mohan, N.; Cummings Bende, E.M.; Maloney, A.J.; Nieves, M.; Sangerman, J.; Lu, A.E.; Hong, M.S.; et al. Automated outlier detection and estimation of missing data. *Comput. Chem. Eng.* **2024**, *180*, 108448. [CrossRef]
588. Camacho, J.; Pérez-Villegas, A.; Rodríguez-Gómez, R.A.; Jiménez-Mañas, E. Multivariate Exploratory Data Analysis (MEDA) Toolbox for Matlab. *Chemom. Intell. Lab. Syst.* **2015**, *143*, 49–57. [CrossRef]
589. Jiang, Y.; Yin, S.; Yang, Y. Comparison of KPI related fault detection algorithms using a newly developed MATLAB toolbox: DB-KIT. In Proceedings of the IECON 2016—42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, Italy, 23–26 October 2016; pp. 7149–7154. [CrossRef]
590. Jiang, Y.; Yin, S. Recent Advances in Key-Performance-Indicator Oriented Prognosis and Diagnosis with a MATLAB Toolbox: DB-KIT. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2849–2858. [CrossRef]
591. González-Martínez, J.; Camacho, J.; Ferrer, A. MVBatch: A matlab toolbox for batch process modeling and monitoring. *Chemom. Intell. Lab. Syst.* **2018**, *183*, 122–133. [CrossRef]
592. Yi, G.; Herdsman, C.; Morris, J. A MATLAB toolbox for data pre-processing and multivariate statistical process control. *Chemom. Intell. Lab. Syst.* **2019**, *194*, 103863. [CrossRef]
593. Villalba, P.; Sanchis, J.; Ferrer, A. A graphical user interface for PCA-based MSPC: A benchmark software for multivariate statistical process control in MATLAB. *Chemom. Intell. Lab. Syst.* **2019**, *185*, 135–152. [CrossRef]
594. Schaeffer, J.; Braatz, R.D. Latent Variable Method Demonstrator—Software for understanding multivariate data analytics algorithms. *Comput. Chem. Eng.* **2022**, *167*, 108014. [CrossRef]
595. Sun, B.; Wang, J.; Hou, B.; Xu, S.; Zhang, K.; He, Z. A data-driven fault detection toolbox based on MATLAB GUIDE. In Proceedings of the 2017 Chinese Automation Congress (CAC), Jinan, China, 20–22 October 2017; pp. 4326–4331. [CrossRef]
596. Alizadeh, E.; Koujok, M.E.; Ragab, A.; Amazouz, M. A Data-Driven Causality Analysis Tool for Fault Diagnosis in Industrial Processes. *IFAC-PapersOnLine* **2018**, *51*, 147–152. [CrossRef]
597. Fleming, P.J.; Wallace, J.J. How not to lie with statistics: The correct way to summarize benchmark results. *Commun. ACM* **1986**, *29*, 218–221. [CrossRef]
598. Bajpai, R.K.; Reub, M. A mechanistic model for penicillin production. *J. Chem. Technol. Biotechnol.* **1980**, *30*, 332–344. [CrossRef]
599. Birol, G.; Ündey, C.; Çinar, A. A modular simulation package for fed-batch fermentation: Penicillin production. *Comput. Chem. Eng.* **2002**, *26*, 1553–1565. [CrossRef]
600. Bartys, M.; Patton, R.; Syfert, M.; de las Heras, S.; Quevedo, J. Introduction to the DAMADICS actuator FDI benchmark study. *Control Eng. Pract.* **2006**, *14*, 577–596. [CrossRef]
601. Alex, J.; Benedetti, L.; Copp, J.; Gernaey, K.V.; Jeppsson, U.; Nopens, I.; Pons, M.N.; Rieger, L.; Rosen, C.; Steyer, J.P.; et al. *Benchmark Simulation Model No. 1 (BSM1)*; Technical Report, IWA Taskgroup on Benchmarking of Control Strategies for WWTPs; Lund University: Lund, Sweden, 2008.

602. Goldrick, S.; Duran-Villalobos, C.A.; Jankauskas, K.; Lovett, D.; Farid, S.S.; Lennox, B. Modern day monitoring and control challenges outlined on an industrial-scale benchmark fermentation process. *Comput. Chem. Eng.* **2019**, *130*, 106471. [CrossRef]
603. Goldrick, S.; Stefan, A.; Lovett, D.; Montague, G.; Lennox, B. The development of an industrial-scale fed-batch fermentation simulation. *J. Biotechnol.* **2015**, *193*, 70–82. [CrossRef] [PubMed]
604. Paul, G.C.; Thomas, C.R. A structured model for hyphal differentiation and penicillin production using Penicillium chrysogenum. *Biotechnol. Bioeng.* **1996**, *51*, 558–572. [CrossRef]
605. Ruiz-Cárcel, C.; Cao, Y.; Mba, D.; Lao, L.; Samuel, R. Statistical process monitoring of a multiphase flow facility. *Control Eng. Pract.* **2015**, *42*, 74–88. [CrossRef]
606. Stief, A.; Tan, R.; Cao, Y.; Ottewill, J.R.; Thornhill, N.F.; Baranowski, J. A heterogeneous benchmark dataset for data analytics: Multiphase flow facility case study. *J. Process Control* **2019**, *79*, 41–55. [CrossRef]
607. Stief, A.; Tan, R.; Cao, Y.; Ottewill, J.R. PRONTO Heterogeneous Benchmark Dataset. Available online: https://zenodo.org/records/1341583 (accessed on 16 January 2024).
608. Vargas, R.E.V.; Munaro, C.J.; Ciarelli, P.M.; Medeiros, A.G.; Amaral, B.G.d.; Barrionuevo, D.C.; Araújo, J.C.D.d.; Ribeiro, J.L.; Magalhães, L.P. A realistic and public dataset with rare undesirable real events in oil wells. *J. Pet. Sci. Eng.* **2019**, *181*, 106223. [CrossRef]
609. Fernandes, N.C.P.; Romanenko, A.; Reis, M.S. Mechanistic Modeling and Simulation for Process Data Generation. *Ind. Eng. Chem. Res.* **2019**, *58*, 17871–17884. [CrossRef]
610. Santander, O.; Kuppuraj, V.; Harrison, C.A.; Baldea, M. An open source fluid catalytic cracker-fractionator model to support the development and benchmarking of process control, machine learning and operation strategies. *Comput. Chem. Eng.* **2022**, *164*, 107900. [CrossRef]
611. Yousef, I.; Rippon, L.D.; Prévost, C.; Shah, S.L.; Gopaluni, R.B. The arc loss challenge: A novel industrial benchmark for process analytics and machine learning. *J. Process Control* **2023**, *128*, 103023. [CrossRef]
612. Wise, B.; Veltkamp, D.; Ricker, N.; Kowalski, B. Application of Multivariate Statistical Process Control (MSPC) to the West Valley Slurry-Fed Ceramic Melter Process. In Proceedings of the Waste Management '91 Proceedings, Tucson, Arizona, 24–28 February 1991.
613. Slama, C.F. Multivariate Statistical Analysis of Data from an Industrial Fluidized Catalytic Cracking Process Using PCA and PLS. Master's Thesis, McMaster University, Hamilton, ON, Canada, 1991.
614. Nomikos, P.; MacGregor, J.F. Multivariate SPC Charts for Monitoring Batch Processes. *Technometrics* **1995**, *37*, 41–59. [CrossRef]
615. Piovoso, M.J.; Kosanovich, K.A.; Pearson, R.K. Monitoring Process Performance in Real-Time. In Proceedings of the 1992 American Control Conference, Chicago, IL, USA, 24–26 June 1992; pp. 2359–2363. [CrossRef]
616. Piovoso, M.J.; Kosanovich, K.A. Applications of multivariate statistical methods to process monitoring and controller design. *Int. J. Control* **1994**, *59*, 743–765. [CrossRef]
617. Kourti, T.; Lee, J.; Macgregor, J.F. Experiences with industrial applications of projection methods for multivariate statistical process control. *Comput. Chem. Eng.* **1996**, *20*, S745–S750. [CrossRef]
618. Ignova, M.; Glassey, J.; Ward, A.; Montague, G. Multivariate statistical methods in bioprocess fault detection and performance forecasting. *Trans. Inst. Meas. Control* **1997**, *19*, 271–279. [CrossRef]
619. Montague, G.; Hiden, H.; Kornfeld, G. Multivariate Statistical Monitoring Procedures for Fermentation Supervision: An Industrial Case Study. *IFAC Proc. Vol.* **1998**, *31*, 399–404. [CrossRef]
620. Lennox, B.; Montague, G.; Hiden, H.; Kornfeld, G.; Goulding, P. Process monitoring of an industrial fed-batch fermentation. *Biotechnol. Bioeng.* **2001**, *74*, 125–135. [CrossRef]
621. Gurden, S.; Martin, E.; Morris, A. The introduction of process chemometrics into an industrial pilot plant laboratory. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 319–330. [CrossRef]
622. Taylor, A. The Application of Principal Component Analysis for Predicting Blast Furnace Stability. *IFAC Proc. Vol.* **1998**, *31*, 223–226. [CrossRef]
623. Neogi, D.; Schlags, C.E. Multivariate Statistical Analysis of an Emulsion Batch Process. *Ind. Eng. Chem. Res.* **1998**, *37*, 3971–3979. [CrossRef]
624. Martin, E.; Morris, A.; Kiparissides, C. Multivariate Statistical Process Control and Process Performance Monitoring. *IFAC Proc. Vol.* **1998**, *31*, 347–356. [CrossRef]
625. Bissessur, Y.; Martin, E.; Morris, A. Monitoring the performance of the paper making process. *Control Eng. Pract.* **1999**, *7*, 1357–1368. [CrossRef]
626. Dudzic, M.; Vaculik, V.; Miletic, I. On-line applications of multivariate statistics at Dofasco. *IFAC Proc. Vol.* **2000**, *33*, 425–430. [CrossRef]
627. Albert, S.; Kinley, R.D. Multivariate statistical monitoring of batch processes: An industrial case study of fermentation supervision. *Trends Biotechnol.* **2001**, *19*, 53–62. [CrossRef] [PubMed]
628. Karim, M.; Hodge, D.; Simon, L. Data-Based Modeling and Analysis of Bioprocesses: Some Real Experiences. *Biotechnol. Prog.* **2003**, *19*, 1591–1605. [CrossRef] [PubMed]
629. Kumar, V.; Sundararaj, U.; Shah, S.L.; Hair, D.; Vande Griend, L.J. Multivariate Statistical Monitoring of a High-Pressure Polymerization Process. *Polym. React. Eng.* **2003**, *11*, 1017–1052. [CrossRef]

630. Kourti, T. Abnormal situation detection and projection methods—Industrial applications. October 28–29, 2003. Hamilton, Ontario, Canada. *Chemom. Intell. Lab. Syst.* **2005**, *76*, 215–220. [CrossRef]
631. Miletic, I.; Quinn, S.; Dudzic, M.; Vaculik, V.; Champagne, M. An industrial perspective on implementing on-line applications of multivariate statistics. *J. Process Control* **2004**, *14*, 821–836. [CrossRef]
632. Kourti, T. Application of latent variable methods to process control and multivariate statistical process control in industry. *Int. J. Adapt. Control Signal Process.* **2005**, *19*, 213–246. [CrossRef]
633. Zhang, Y.; Dudzic, M.S. Online monitoring of steel casting processes using multivariate statistical technologies: From continuous to transitional operations. *J. Process Control* **2006**, *16*, 819–829. [CrossRef]
634. Qin, S.J.; Cherry, G.; Good, R.; Wang, J.; Harrison, C.A. Semiconductor manufacturing process control and monitoring: A fab-wide framework. *J. Process Control* **2006**, *16*, 179–191. [CrossRef]
635. Chiang, L.H.; Colegrove, L.F. Industrial implementation of on-line multivariate quality control. *Chemom. Intell. Lab. Syst.* **2007**, *88*, 143–153. [CrossRef]
636. Kano, M.; Nakagawa, Y. Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry. *Comput. Chem. Eng.* **2008**, *32*, 12–24. [CrossRef]
637. Miletic, I.; Boudreau, F.; Dudzic, M.; Kotuza, G.; Ronholm, L.; Vaculik, V.; Zhang, Y. Experiences in applying data-driven modelling technology to steelmaking processes. *Can. J. Chem. Eng.* **2008**, *86*, 937–946. [CrossRef]
638. AlGhazzawi, A.; Lennox, B. Monitoring a complex refining process using multivariate statistics. *Control Eng. Pract.* **2008**, *16*, 294–307. [CrossRef]
639. Vanhatalo, E. Multivariate process monitoring of an experimental blast furnace. *Qual. Reliab. Eng. Int.* **2010**, *26*, 495–508. [CrossRef]
640. Darkow, T.; Dittmar, R.; Timm, H. Real-time application of multivariate statistical methods for early event detection in an industrial slurry stripper. *IFAC Proc. Vol.* **2014**, *47*, 8879–8884. [CrossRef]
641. Dumarey, M.; Hermanto, M.; Airiau, C.; Shapland, P.; Robinson, H.; Hamilton, P.; Berry, M. Advances in Continuous Active Pharmaceutical Ingredient (API) Manufacturing: Real-time Monitoring Using Multivariate Tools. *J. Pharm. Innov.* **2018**, 14, 359–372. [CrossRef]
642. Patwardhan, R.S.; Hamadah, H.A.; Patel, K.M.; Hafiz, R.H.; Al-Gwaiz, M.M. Applications of Advanced Analytics at Saudi Aramco: A Practitioners' Perspective. *Ind. Eng. Chem. Res.* **2019**, *58*, 11338–11351. [CrossRef]
643. Klanderman, M.C.; Newhart, K.B.; Cath, T.Y.; Hering, A.S. Case studies in real-time fault isolation in a decentralized wastewater treatment facility. *J. Water Process Eng.* **2020**, *38*, 101556. [CrossRef]
644. Kumar, A.; Bhattacharya, A.; Flores-Cerrillo, J. Data-driven process monitoring and fault analysis of reformer units in hydrogen plants: Industrial application and perspectives. *Comput. Chem. Eng.* **2020**, *136*, 106756. [CrossRef]
645. Sivalingam, G.; Soni, N.J.; Vakil, S.M. Detection of decomposition for high pressure ethylene/vinyl acetate copolymerization in autoclave reactor using principal component analysis on heat balance model. *Can. J. Chem. Eng.* **2015**, *93*, 1063–1075. [CrossRef]
646. Harrou, F.; Dairi, A.; Sun, Y.; Senouci, M. Statistical monitoring of a wastewater treatment plant: A case study. *J. Environ. Manag.* **2018**, *223*, 807–814. [CrossRef]
647. Sad, C.M.; da Silva, M.; dos Santos, F.D.; Pereira, L.B.; Corona, R.R.; Silva, S.R.; Portela, N.A.; Castro, E.V.; Filgueiras, P.R.; Lacerda, V. Multivariate data analysis applied in the evaluation of crude oil blends. *Fuel* **2019**, *239*, 421–428. [CrossRef]
648. Clavijo, N.; Melo, A.; Câmara, M.M.; Feital, T.; Anzai, T.K.; Diehl, F.C.; Thompson, P.H.; Pinto, J.C. Development and Application of a Data-Driven System for Sensor Fault Diagnosis in an Oil Processing Plant. *Processes* **2019**, *7*, 436. [CrossRef]
649. Nachtergaele, P.; Thybaut, J.W.; De Meester, S.; Drijvers, D.; Saeys, W.; Dewulf, J. Multivariate Analysis of Industrial Biorefinery Processes: Strategy for Improved Process Understanding with Case Studies in Fatty Acid Production. *Ind. Eng. Chem. Res.* **2020**, *59*, 7732–7745. [CrossRef]
650. Li, X.; Xue, F.; Qin, L.; Zhou, K.; Chen, Z.; Ge, Z.; Chen, X.; Song, K. A recursively updated Map-Reduce based PCA for monitoring the time-varying fluorochemical engineering processes with big data. *Chemom. Intell. Lab. Syst.* **2020**, *206*, 104167. [CrossRef]
651. Rippon, L.D.; Yousef, I.; Hosseini, B.; Bouchoucha, A.; Beaulieu, J.F.; Prévost, C.; Ruel, M.; Shah, S.L.; Gopaluni, R.B. Representation Learning and Predictive Classification: Application with an Electric Arc Furnace. *Comput. Chem. Eng.* **2021**, *150*, 107304. [CrossRef]
652. Tian, W.; Wang, S.; Sun, S.; Li, C.; Lin, Y. Intelligent prediction and early warning of abnormal conditions for fluid catalytic cracking process. *Chem. Eng. Res. Des.* **2022**, *181*, 304–320. [CrossRef]
653. Nadir, F.; Messaoud, B.; Elias, H. Utilizing Principal Component Analysis for the Identification of Gas Turbine Defects. *J. Fail. Anal. Prev.* **2023**. [CrossRef]
654. Bi, X.; Qin, R.; Wu, D.; Zheng, S.; Zhao, J. One step forward for smart chemical process fault detection and diagnosis. *Comput. Chem. Eng.* **2022**, *164*, 107884. [CrossRef]
655. Okada, K.F.A.; Silva de Morais, A.; Oliveira-Lopes, L.C.; Ribeiro, L. A Survey on Fault Detection and Diagnosis Methods. In Proceedings of the 2021 14th IEEE International Conference on Industry Applications (INDUSCON), São Paulo, Brazil, 15–18 August 2021; pp. 1422–1429. [CrossRef]
656. Park, Y.J.; Fan, S.K.S.; Hsu, C.Y. A Review on Fault Detection and Process Diagnostics in Industrial Processes. *Processes* **2020**, *8*, 1123. [CrossRef]
657. Diez-Olivan, A.; Del Ser, J.; Galar, D.; Sierra, B. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Inf. Fusion* **2019**, *50*, 92–111. [CrossRef]

658. Dai, X.; Gao, Z. From Model, Signal to Knowledge: A Data-Driven Perspective of Fault Detection and Diagnosis. *IEEE Trans. Ind. Inform.* **2013**, *9*, 2226–2238. [CrossRef]
659. Bradley, W.; Kim, J.; Kilwein, Z.; Blakely, L.; Eydenberg, M.; Jalvin, J.; Laird, C.; Boukouvala, F. Perspectives on the integration between first-principles and data-driven modeling. *Comput. Chem. Eng.* **2022**, *166*, 107898. [CrossRef]
660. Sansana, J.; Joswiak, M.N.; Castillo, I.; Wang, Z.; Rendall, R.; Chiang, L.H.; Reis, M.S. Recent trends on hybrid modeling for Industry 4.0. *Comput. Chem. Eng.* **2021**, *151*, 107365. [CrossRef]
661. Rajulapati, L.; Chinta, S.; Shyamala, B.; Rengaswamy, R. Integration of machine learning and first principles models. *AIChE J.* **2022**, *68*, e17715. [CrossRef]
662. Yan, W.; Wang, J.; Lu, S.; Zhou, M.; Peng, X. A Review of Real-Time Fault Diagnosis Methods for Industrial Smart Manufacturing. *Processes* **2023**, *11*, 369. [CrossRef]
663. Perera, Y.S.; Ratnaweera, D.A.A.C.; Dasanayaka, C.H.; Abeykoon, C. The role of artificial intelligence-driven soft sensors in advanced sustainable process industries: A critical review. *Eng. Appl. Artif. Intell.* **2023**, *121*, 105988. [CrossRef]
664. Jiang, Y.; Yin, S.; Dong, J.; Kaynak, O. A Review on Soft Sensors for Monitoring, Control and Optimization of Industrial Processes. *IEEE Sens. J.* **2020**, *21*, 12868–12881. [CrossRef]
665. Qin, S.J.; Dong, Y.; Zhu, Q.; Wang, J.; Liu, Q. Bridging systems theory and data science: A unifying review of dynamic latent variable analytics and process monitoring. *Annu. Rev. Control* **2020**, *50*, 29–48. [CrossRef]
666. Liu, J.; Hou, L.; Wang, X.; Zhang, R.; Sun, X.; Xu, L.; Yu, Q. Explainable fault diagnosis of gas-liquid separator based on fully convolutional neural network. *Comput. Chem. Eng.* **2021**, *155*, 107535. [CrossRef]
667. Kim, M.S.; Yun, J.P.; Park, P. An Explainable Convolutional Neural Network for Fault Diagnosis in Linear Motion Guide. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4036–4045. [CrossRef]
668. Agarwal, P.; Tamer, M.; Budman, H. Explainability: Relevance based Dynamic Deep Learning Algorithm for Fault Detection and Diagnosis in Chemical Processes. *Comput. Chem. Eng.* **2021**, *154*, 107467. [CrossRef]
669. Bhakte, A.; Pakkiriswamy, V.; Srinivasan, R. An explainable artificial intelligence based approach for interpretation of fault classification results from deep neural networks. *Chem. Eng. Sci.* **2022**, *250*, 117373. [CrossRef]
670. Choi, S.H.; Lee, J.M. Explainable Fault Diagnosis Model using Stacked Autoencoder and Kernel SHAP. In Proceedings of the 2022 IEEE International Symposium on Advanced Control of Industrial Processes (AdCONIP), Vancouver, BC, Canada, 7–9 August 2022; pp. 182–187. [CrossRef]
671. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In Proceedings of the the KDD '16: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144. [CrossRef]
672. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the the NIPS'17: 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777.
673. Kaur, H.; Nori, H.; Jenkins, S.; Caruana, R.; Wallach, H.; Wortman Vaughan, J. Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning. In Proceedings of the CHI '20: 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–14. [CrossRef]
674. Rudin, C.; Chen, C.; Chen, Z.; Huang, H.; Semenova, L.; Zhong, C. Interpretable machine learning: Fundamental principles and 10 grand challenges. *Stat. Surv.* **2022**, *16*, 1–85. [CrossRef]
675. Burkart, N.; Huber, M.F. A Survey on the Explainability of Supervised Machine Learning. *J. Artif. Intell. Res.* **2021**, *70*, 245–317. [CrossRef]
676. Molnar, C.; Casalicchio, G.; Bischl, B. Interpretable Machine Learning—A Brief History, State-of-the-Art and Challenges. In Proceedings of the ECML PKDD 2020 Workshops, Ghent, Belgium, 14–18 September 2020; Communications in Computer and Information Science; Koprinska, I., Kamp, M., Appice, A., Loglisci, C., Antonie, L., Zimmermann, A., Guidotti, R., Özgöbek, O., Ribeiro, R.P., Gavaldà, R., et al., Eds.; Springer: Cham, Switzerland, 2020; pp. 417–431. [CrossRef]
677. Linardatos, P.; Papastefanopoulos, V.; Kotsiantis, S. Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy* **2021**, *23*, 18. [CrossRef]
678. Carvalho, D.V.; Pereira, E.M.; Cardoso, J.S. Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics* **2019**, *8*, 832. [CrossRef]
679. Carter, A.; Imtiaz, S.; Naterer, G.F. Review of interpretable machine learning for process industries. *Process Saf. Environ. Prot.* **2023**, *170*, 647–659. [CrossRef]
680. Harinarayan, R.R.A.; Shalinie, S.M. XFDDC: eXplainable Fault Detection Diagnosis and Correction framework for chemical process systems. *Process Saf. Environ. Prot.* **2022**, *165*, 463–474. [CrossRef]
681. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the KDD '16: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]
682. Bhakte, A.; Chakane, M.; Srinivasan, R. Alarm-based explanations of process monitoring results from deep neural networks. *Comput. Chem. Eng.* **2023**, *179*, 108442. [CrossRef]
683. Ye, L.; Wu, H.; Chen, Y.; Fei, Z. Interpret what a Convolutional Neural Network learns for fault detection and diagnosis in process systems. *J. Process Control* **2023**, *131*, 103086. [CrossRef]
684. Chollet, F. *Deep Learning with Python*, 2nd ed.; Manning Publications: Shelter Island, NY, USA, 2021.

685. Stevens, E.; Antiga, L.; Viehmann, T. *Deep Learning with PyTorch: Build, Train, and Tune Neural Networks Using Python Tools*, 1st ed.; Manning: Shelter Island, NY, USA, 2020.
686. Wang, Y.; Hill, K.J.; Foley, E.C. Computer programming with Python for industrial and systems engineers: Perspectives from an instructor and students. *Comput. Appl. Eng. Educ.* **2017**, *25*, 800–811. [CrossRef]
687. Ahmad, S.; Purdy, S. Real-Time Anomaly Detection for Streaming Analytics. *arXiv* **2016**, arXiv:1607.02480.
688. Ahmad, S.; Lavin, A.; Purdy, S.; Agha, Z. Unsupervised real-time anomaly detection for streaming data. *Neurocomputing* **2017**, *262*, 134–147. [CrossRef]
689. Akcay, S.; Ameln, D.; Vaidya, A.; Lakshmanan, B.; Ahuja, N.; Genc, U. Anomalib: A Deep Learning Library for Anomaly Detection. In Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 16–19 October 2022; pp. 1706–1710. [CrossRef]
690. Barbariol, T.; Chiara, F.D.; Marcato, D.; Susto, G.A. A Review of Tree-Based Approaches for Anomaly Detection. In *Control Charts and Machine Learning for Anomaly Detection in Manufacturing*; Tran, K.P., Ed.; Springer Series in Reliability Engineering; Springer International Publishing: Cham, Switerland, 2022; pp. 149–185. [CrossRef]
691. Ariyaluran Habeeb, R.A.; Nasaruddin, F.; Gani, A.; Amanullah, M.A.; Abaker Targio Hashem, I.; Ahmed, E.; Imran, M. Clustering-based real-time anomaly detection—A breakthrough in big data technologies. *Trans. Emerg. Telecommun. Technol.* **2022**, *33*, e3647. [CrossRef]
692. Blázquez-García, A.; Conde, A.; Mori, U.; Lozano, J.A. A Review on Outlier/Anomaly Detection in Time Series Data. *ACM Comput. Surv.* **2022**, *54*, 1–33. [CrossRef]
693. Andrade, T.; Gama, J.; Ribeiro, R.P.; Sousa, W.; Carvalho, A. Anomaly Detection in Sequential Data: Principles and Case Studies. In *Wiley Encyclopedia of Electrical and Electronics Engineering*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2019; pp. 1–14. [CrossRef]
694. Carletti, M.; Terzi, M.; Susto, G.A. Interpretable Anomaly Detection with DIFFI: Depth-based feature importance of Isolation Forest. *Eng. Appl. Artif. Intell.* **2023**, *119*, 105730. [CrossRef]
695. Chalapathy, R.; Chawla, S. Deep Learning for Anomaly Detection: A Survey. *arXiv* **2019**, arXiv:1901.03407.
696. Foorthuis, R. On the nature and types of anomalies: A review of deviations in data. *Int. J. Data Sci. Anal.* **2021**, *12*, 297–331. [CrossRef]
697. Garg, A.; Zhang, W.; Samaran, J.; Savitha, R.; Foo, C.S. An Evaluation of Anomaly Detection and Diagnosis in Multivariate Time Series. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 2508–2517. [CrossRef]
698. Goldstein, M.; Uchida, S. A Comparative Evaluation of Unsupervised Anomaly Detection Algorithms for Multivariate Data. *PLoS ONE* **2016**, *11*, e0152173. [CrossRef]
699. Han, S.; Hu, X.; Huang, H.; Jiang, M.; Zhao, Y. ADBench: Anomaly Detection Benchmark. In *Advances in Neural Information Processing Systems, Proceedings of the NeurIPS 2022, New Orleans, LA, USA, 28 November–9 December 2022*; Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., Oh, A., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2022; Volume 35, pp. 32142–32159.
700. Hilal, W.; Gadsden, S.A.; Yawney, J. Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. *Expert Syst. Appl. Int. J.* **2022**, *193*, 116429. [CrossRef]
701. Lakshminarayan, C.; Alvarado, A.S.; Principe, J.C.; Kriminger, E. Anomaly Detection in Streaming Data. U.S. Patent US9218527B2, 22 December 2015.
702. Lavin, A.; Ahmad, S. Evaluating Real-Time Anomaly Detection Algorithms—The Numenta Anomaly Benchmark. In Proceedings of the 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 9–11 December 2015; pp. 38–44. [CrossRef]
703. Ruff, L.; Kauffmann, J.R.; Vandermeulen, R.A.; Montavon, G.; Samek, W.; Kloft, M.; Dietterich, T.G.; Muller, K.R. A Unifying Review of Deep and Shallow Anomaly Detection. *Proc. IEEE* **2021**, *109*, 756–795. [CrossRef]
704. Schmidl, S.; Wenig, P.; Papenbrock, T. Anomaly detection in time series: A comprehensive evaluation. *Proc. VLDB Endow.* **2022**, *15*, 1779–1797. [CrossRef]
705. Sgueglia, A.; Di Sorbo, A.; Visaggio, C.A.; Canfora, G. A systematic literature review of IoT time series anomaly detection solutions. *Future Gener. Comput. Syst.* **2022**, *134*, 170–186. [CrossRef]
706. Spies, N.C.; Farnsworth, C.W.; Jackups, R. Data-Driven Anomaly Detection in Laboratory Medicine: Past, Present, and Future. *J. Appl. Lab. Med.* **2023**, *8*, 162–179. [CrossRef] [PubMed]
707. Talagala, P.D.; Hyndman, R.J.; Smith-Miles, K.; Kandanaarachchi, S.; Muñoz, M.A. Anomaly Detection in Streaming Nonstationary Temporal Data. *J. Comput. Graph. Stat.* **2019**, *29*, 13–27. [CrossRef]
708. Ul Hassan, M.; Rehmani, M.H.; Chen, J. Anomaly Detection in Blockchain Networks: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 289–318. [CrossRef]
709. Xia, X.; Pan, X.; Li, N.; He, X.; Ma, L.; Zhang, X.; Ding, N. GAN-based anomaly detection: A review. *Neurocomputing* **2022**, *493*, 497–535. [CrossRef]
710. Bouguelia, M.R.; Nowaczyk, S.; Payberah, A.H. An adaptive algorithm for anomaly and novelty detection in evolving data streams. *Data Min. Knowl. Discov.* **2018**, *32*, 1597–1633. [CrossRef]
711. Cejnek, M.; Bukovsky, I. Concept drift robust adaptive novelty detection for data streams. *Neurocomputing* **2018**, *309*, 46–53. [CrossRef]
712. Clifton, D.A.; Hugueny, S.; Tarassenko, L. Novelty Detection with Multivariate Extreme Value Statistics. *J. Signal Process. Syst.* **2011**, *65*, 371–389. [CrossRef]

713. Decke, J.; Schmeißing, J.; Botache, D.; Bieshaar, M.; Sick, B.; Gruhl, C. NDNET: A Unified Framework for Anomaly and Novelty Detection. In Proceedings of the Architecture of Computing Systems, Heilbronn, Germany, 13–15 September 2022; Lecture Notes in Computer Science; Schulz, M., Trinitis, C., Papadopoulou, N., Pionteck, T., Eds.; Springer: Cham, Switzerland, 2022; pp. 197–210. [CrossRef]
714. Ding, X.; Li, Y.; Belatreche, A.; Maguire, L.P. An experimental evaluation of novelty detection methods. *Neurocomputing* **2014**, *135*, 313–327. [CrossRef]
715. Domingues, R.; Michiardi, P.; Barlet, J.; Filippone, M. A comparative evaluation of novelty detection algorithms for discrete sequences. *Artif. Intell. Rev.* **2020**, *53*, 3787–3812. [CrossRef]
716. Faria, E.R.; Goncalves, I.R.; Gama, J.a.; Carvalho, A.C.P.d.L.F. Evaluation of Multiclass Novelty Detection Algorithms for Data Streams. *IEEE Trans. Knowl. Data Eng.* **2015**, *27*, 2961–2973. [CrossRef]
717. Ghosal, T.; Saikh, T.; Biswas, T.; Ekbal, A.; Bhattacharyya, P. Novelty Detection: A Perspective from Natural Language Processing. *Comput. Linguist.* **2022**, *48*, 77–117. [CrossRef]
718. Gruhl, C.M. *Novelty Detection for Multivariate Data Streams with Probalistic Models*; Kassel University Press: Kassel, Germany, 2022. [CrossRef]
719. La Grassa, R.; Gallo, I.; Landro, N. OCmst: One-class novelty detection using convolutional neural network and minimum spanning trees. *Pattern Recognit. Lett.* **2022**, *155*, 114–120. [CrossRef]
720. Markou, M.; Singh, S. Novelty detection: A review—Part 1: Statistical approaches. *Signal Process.* **2003**, *83*, 2481–2497. [CrossRef]
721. Markou, M.; Singh, S. Novelty detection: A review—Part 2: Neural network based approaches. *Signal Process.* **2003**, *83*, 2499–2521. [CrossRef]
722. Pimentel, M.A.; Clifton, D.A.; Clifton, L.; Tarassenko, L. A review of novelty detection. *Signal Process.* **2014**, *99*, 215–249. [CrossRef]
723. Skvara, V.; Pevny, T.; Smidl, V. Are generative deep models for novelty detection truly better? *arXiv* **2018**, arXiv:1807.05027.
724. Tarassenko, L.; Clifton, D.A.; Bannister, P.R.; King, S.; King, D. Novelty Detection. In *Encyclopedia of Structural Health Monitoring*; Boller, C., Chang, F.K., Fujino, Y., Eds.; John Wiley: Chichester, West Sussex, UK, 2009.
725. Adams, R.P.; MacKay, D.J.C. Bayesian Online Changepoint Detection. *arXiv* **2007**, arXiv:0710.3742.
726. Alippi, C.; Boracchi, G.; Carrera, D.; Roveri, M. Change Detection in Multivariate Datastreams: Likelihood and Detectability Loss. *arXiv* **2015**, arXiv:1510.04850.
727. Aminikhanghahi, S.; Cook, D.J. A survey of methods for time series change point detection. *Knowl. Inf. Syst.* **2017**, *51*, 339–367. [CrossRef]
728. Cabrieto, J.; Tuerlinckx, F.; Kuppens, P.; Grassmann, M.; Ceulemans, E. Detecting correlation changes in multivariate time series: A comparison of four non-parametric change point detection methods. *Behav. Res. Methods* **2017**, *49*, 988–1005. [CrossRef]
729. Cabrieto, J.; Tuerlinckx, F.; Kuppens, P.; Wilhelm, F.H.; Liedlgruber, M.; Ceulemans, E. Capturing correlation changes by applying kernel change point detection on the running correlations. *Inf. Sci.* **2018**, *447*, 117–139. [CrossRef]
730. Cabrieto, J.; Tuerlinckx, F.; Kuppens, P.; Hunyadi, B.; Ceulemans, E. Testing for the Presence of Correlation Changes in a Multivariate Time Series: A Permutation Based Approach. *Sci. Rep.* **2018**, *8*, 769. [CrossRef] [PubMed]
731. Chen, Y.; Wang, T.; Samworth, R.J. High-Dimensional, Multiscale Online Changepoint Detection. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2022**, *84*, 234–266. [CrossRef]
732. Gupta, M.; Wadhvani, R.; Rasool, A. Real-time Change-Point Detection: A deep neural network-based adaptive approach for detecting changes in multivariate time series data. *Expert Syst. Appl.* **2022**, *209*, 118260. [CrossRef]
733. Kuncheva, L.I. Change Detection in Streaming Multivariate Data Using Likelihood Detectors. *IEEE Trans. Knowl. Data Eng.* **2013**, *25*, 1175–1180. [CrossRef]
734. Niu, Y.S.; Hao, N.; Zhang, H. Multiple Change-Point Detection: A Selective Overview. *Stat. Sci.* **2016**, *31*, 611–623. [CrossRef]
735. Preuss, P.; Puchstein, R.; Dette, H. Detection of Multiple Structural Breaks in Multivariate Time Series. *J. Am. Stat. Assoc.* **2015**, *110*, 654–668. [CrossRef]
736. Staudacher, M.; Telser, S.; Amann, A.; Hinterhuber, H.; Ritsch-Marte, M. A new method for change-point detection developed for on-line analysis of the heart beat variability during sleep. *Phys. A Stat. Mech. Its Appl.* **2005**, *349*, 582–596. [CrossRef]
737. Tran, D. Automated Change Detection and Reactive Clustering in Multivariate Streaming Data. In Proceedings of the 2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF), Danang, Vietnam, 20–22 March 2019. [CrossRef]
738. Truong, C.; Oudre, L.; Vayatis, N. Selective review of offline change point detection methods. *Signal Process.* **2020**, *167*, 107299. [CrossRef]
739. Xie, Y.; Huang, J.; Willett, R. Change-Point Detection for High-Dimensional Time Series With Missing Data. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 12–27. [CrossRef]
740. Xie, L.; Zou, S.; Xie, Y.; Veeravalli, V.V. Sequential (Quickest) Change Detection: Classical Results and New Directions. *IEEE J. Sel. Areas Inf. Theory* **2021**, *2*, 494–514. [CrossRef]
741. Xuan, X.; Murphy, K. Modeling changing dependency structure in multivariate time series. In Proceedings of the the 24th International Conference on Machine Learning—ICML '07, Corvalis, OR, USA, 20–24 June 2007; pp. 1055–1062. [CrossRef]
742. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [CrossRef]

743. Saini, K.; Dhami, S.S.; Vanraj. Predictive Monitoring of Incipient Faults in Rotating Machinery: A Systematic Review from Data Acquisition to Artificial Intelligence. *Arch. Comput. Methods Eng.* **2022**, *29*, 4005–4026. [CrossRef]
744. Fernandes, M.; Corchado, J.M.; Marreiros, G. Machine learning techniques applied to mechanical fault diagnosis and fault prognosis in the context of real industrial manufacturing use-cases: A systematic literature review. *Appl. Intell.* **2022**, *52*, 14246–14280. [CrossRef]
745. Ma, J.; Jiang, J. Applications of fault detection and diagnosis methods in nuclear power plants: A review. *Prog. Nucl. Energy* **2011**, *53*, 255–266. [CrossRef]
746. Katser, I.D.; Kozitsin, V.O.; Maksimov, I.V.; Larionov, D.A.; Kotsoev, K.I. Data pre-processing methods for NPP equipment diagnostics algorithms: An overview. *Nucl. Energy Technol.* **2021**, *7*, 111–125. [CrossRef]
747. Khoukhi, A.; Khalid, M.H. Hybrid computing techniques for fault detection and isolation, a review. *Comput. Electr. Eng.* **2015**, *43*, 17–32. [CrossRef]
748. Abbasi, A.R. Fault detection and diagnosis in power transformers: A comprehensive review and classification of publications and methods. *Electr. Power Syst. Res.* **2022**, *209*, 107990. [CrossRef]
749. Abid, A.; Khan, M.T.; Iqbal, J. A review on fault detection and diagnosis techniques: Basics and beyond. *Artif. Intell. Rev.* **2020**, *54*, 3639–3664. [CrossRef]
750. Alobaidy, M.; Abdul Jabbar, J.; Al khayyt, S. Faults Diagnosis in Robot Systems: A Review. *Al-Rafidain Eng. J. (AREJ)* **2020**, *25*, 166–177. [CrossRef]
751. Gomes, H.M.; Read, J.; Bifet, A.; Barddal, J.P.; Gama, J. Machine learning for streaming data: State of the art, challenges, and opportunities. *ACM SIGKDD Explor. Newsl.* **2019**, *21*, 6–22. [CrossRef]
752. PhridviRaj, M.; GuruRao, C. Data Mining—Past, Present and Future—A Typical Survey on Data Streams. *Procedia Technol.* **2014**, *12*, 255–263. [CrossRef]
753. Psaltis, A. *Streaming Data: Understanding the Real-Time Pipeline*, 1st ed.; Manning Publications: Shelter Island, NY, USA, 2017.
754. Agrahari, S.; Singh, A.K. Concept Drift Detection in Data Stream Mining: A literature review. *J. King Saud Univ. Comput. Inf. Sci.* **2022**, *34*, 9523–9540. [CrossRef]
755. Janardan; Mehta, S. Concept drift in Streaming Data Classification: Algorithms, Platforms and Issues. *Procedia Comput. Sci.* **2017**, *122*, 804–811. [CrossRef]
756. Zliobaite, I.; Pechenizkiy, M.; Gama, J. An Overview of Concept Drift Applications. In *Big Data Analysis: New Algorithms for a New Society*; Japkowicz, N., Stefanowski, J., Eds.; Springer International Publishing: Cham, Switzerland, 2016; Volume 16, pp. 91–114. [CrossRef]
757. Khamassi, I.; Sayed-Mouchaweh, M.; Hammami, M.; Ghedira, K. Discussion and review on evolving data streams and concept drift adapting. *Evol. Syst.* **2018**, *9*, 1–23. [CrossRef]
758. Souza, V.M.A.; dos Reis, D.M.; Maletzke, A.G.; Batista, G.E.A.P.A. Challenges in benchmarking stream learning algorithms with real-world data. *Data Min. Knowl. Discov.* **2020**, *34*, 1805–1858. [CrossRef]
759. Tsymbal, A. *The Problem of Concept Drift: Definitions and Related Work*; Computer Science Department, Trinity College Dublin: Dublin, Ireland, 2004.
760. Kadlec, P.; Gabrys, B.; Strandt, S. Data-driven Soft Sensors in the process industry. *Comput. Chem. Eng.* **2009**, *33*, 795–814. [CrossRef]
761. Jeong, D.H.; Lee, J.M. Ensemble learning based latent variable model predictive control for batch trajectory tracking under concept drift. *Comput. Chem. Eng.* **2020**, *139*, 106875. [CrossRef]
762. Sun, Z.; Tang, J.; Qiao, J.; Cui, C. Review of Concept Drift Detection Method for Industrial Process Modeling. In Proceedings of the 2020 39th Chinese Control Conference (CCC), Shenyang, China, 27–29 July 2020; pp. 5754–5759. [CrossRef]
763. Barros, R.S.M.d.; Santos, S.G.T.d.C. An overview and comprehensive comparison of ensembles for concept drift. *Inf. Fusion* **2019**, *52*, 213–244. [CrossRef]
764. Cano, A.; Krawczyk, B. Kappa Updated Ensemble for drifting data stream mining. *Mach. Learn.* **2020**, *109*, 175–218. [CrossRef]
765. Krawczyk, B.; Minku, L.L.; Gama, J.; Stefanowski, J.; Wozniak, M. Ensemble learning for data stream analysis: A survey. *Inf. Fusion* **2017**, *37*, 132–156. [CrossRef]
766. Pietruczuk, L.; Rutkowski, L.; Jaworski, M.; Duda, P. How to adjust an ensemble size in stream data mining? *Inf. Sci.* **2017**, *381*, 46–54. [CrossRef]
767. Moshtaghi, M.; Leckie, C.; Bezdek, J.C. Online Clustering of Multivariate Time-series. In Proceedings of the the 2016 SIAM International Conference on Data Mining, Miami, FL, USA, 5–7 May 2016; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2016; pp. 360–368. [CrossRef]
768. Ramirez-Gallego, S.; Krawczyk, B.; García, S.; Wozniak, M.; Herrera, F. A survey on data preprocessing for data stream mining: Current status and future directions. *Neurocomputing* **2017**, *239*, 39–57. [CrossRef]
769. Gupta, M.; Gao, J.; Aggarwal, C.C.; Han, J. Outlier Detection for Temporal Data: A Survey. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 2250–2267. [CrossRef]
770. Sadik, S.; Gruenwald, L. Research issues in outlier detection for data streams. *ACM SIGKDD Explor. Newsl.* **2014**, *15*, 33–40. [CrossRef]
771. Li, Y.; Ge, T.; Chen, C. Data Stream Event Prediction Based on Timing Knowledge and State Transitions. *Proc. VLDB Endow.* **2020**, *13*, 1779–1792. [CrossRef]

772. Chiang, L.H.; Braatz, R.D. Process monitoring using causal map and multivariate statistics: Fault detection and identification. *Chemom. Intell. Lab. Syst.* **2003**, *65*, 159–178. [CrossRef]
773. Hu, J.; Zhang, L.; Cai, Z.; Wang, Y.; Wang, A. Fault propagation behavior study and root cause reasoning with dynamic Bayesian network based framework. *Process Saf. Environ. Prot.* **2015**, *97*, 25–36. [CrossRef]
774. Kathari, S.; Tangirala, A.K. A Novel Framework for Causality Analysis of Deterministic Dynamical Processes. *Ind. Eng. Chem. Res.* **2022**, *61*, 18426–18444. [CrossRef]
775. Li, G.; Qin, S.J.; Yuan, T. Data-driven root cause diagnosis of faults in process industries. *Chemom. Intell. Lab. Syst.* **2016**, *159*, 1–11. [CrossRef]
776. Naghoosi, E.; Huang, B.; Domlan, E.; Kadali, R. Information transfer methods in causality analysis of process variables with an industrial application. *J. Process Control* **2013**, *23*, 1296–1305. [CrossRef]
777. Wang, S.; Zhao, Q.; Han, Y.; Wang, J. Root cause diagnosis for complex industrial process faults via spatiotemporal coalescent based time series prediction and optimized Granger causality. *Chemom. Intell. Lab. Syst.* **2023**, *233*, 104728. [CrossRef]
778. Yang, S.; Bequette, B.W. Observational process data analytics using causal inference. *AIChE J.* **2023**, *69*, e17986. [CrossRef]
779. Yuan, T.; Qin, S.J. Root cause diagnosis of plant-wide oscillations using Granger causality. *IFAC Proc. Vol.* **2012**, *45*, 160–165. [CrossRef]
780. Zope, K.; Singhal, T.; Nistala, S.H.; Runkana, V. Transfer Entropy-Based Automated Fault Traversal and Root Cause Identification in Complex Nonlinear Industrial Processes. *Ind. Eng. Chem. Res.* **2023**, *62*, 4002–4018. [CrossRef]
781. Huang, H.; Wang, R.; Zhou, K.; Ning, L.; Song, K. CausalViT: Domain generalization for chemical engineering process fault detection and diagnosis. *Process Saf. Environ. Prot.* **2023**, *176*, 155–165. [CrossRef]
782. Wang, J.G.; Chen, R.; Ye, X.Y.; Yao, Y.; Xie, Z.T.; Ma, S.W.; Liu, L.L. Data-driven root cause diagnosis of process disturbances by exploring causality change among variables. *J. Process Control* **2023**, *129*, 103062. [CrossRef]
783. Shojaie, A.; Fox, E.B. Granger Causality: A Review and Recent Advances. *Annu. Rev. Stat. Its Appl.* **2022**, *9*, 289–319. [CrossRef] [PubMed]
784. Guo, R.; Cheng, L.; Li, J.; Hahn, P.R.; Liu, H. A Survey of Learning Causality with Data: Problems and Methods. *ACM Comput. Surv.* **2021**, *53*, 1–37. [CrossRef]
785. Clavijo, N.A.; Diehl, F.C.; Gerevini, G.G.; Melo, P.A.; Lemos, T.; Campos, L.F.d.O.; Pinto, J.C. Chaotic dynamics induced by anti-slug control in offshore oil production plants. *J. Pet. Sci. Eng.* **2022**, *215*, 110716. [CrossRef]
786. Clavijo, N.A.; Pinto, J.C.; Melo, P.A. Continuous diabatic free-radical solution polymerization reactors: Search engines for non-linear dynamical solutions. *Chem. Eng. Sci.* **2022**, *248*, 117221. [CrossRef]
787. Melo, P.A.; Biscaia, E.C.; Pinto, J.C. The bifurcation behavior of continuous free-radical solution loop polymerization reactors. *Chem. Eng. Sci.* **2003**, *58*, 2805–2821. [CrossRef]
788. Ourique, C.O.; Biscaia, E.C.; Pinto, J.C. The use of particle swarm optimization for dynamical analysis in chemical processes. *Comput. Chem. Eng.* **2002**, *26*, 1783–1793. [CrossRef]
789. Pinto, J.; Ray, W. The dynamic behavior of continuous solution polymerization reactors—VII. Experimental study of a copolymerization reactor. *Chem. Eng. Sci.* **1995**, *50*, 715–736. [CrossRef]
790. Uppal, A.; Ray, W.; Poore, A. On the dynamic behavior of continuous stirred tank reactors. *Chem. Eng. Sci.* **1974**, *29*, 967–985. [CrossRef]
791. Diks, C. *Nonlinear Time Series Analysis: Methods and Applications*; WSPC: Singapore; River Edge, NJ, USA, 1999.
792. Gooijer, J.G.D. *Elements of Nonlinear Time Series Analysis and Forecasting*; Springer: Cham, Switzerland, 2017.
793. Huffaker, R.; Bittelli, M.; Rosa, R. *Nonlinear Time Series Analysis with R*, 1st ed.; Oxford University Press: Oxford, UK; New York, NY, USA, 2017.
794. Kantz, H.; Schreiber, T. *Nonlinear Time Series Analysis*, 2nd ed.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2004.
795. Dash, S.; Venkatasubramanian, V. Challenges in the industrial applications of fault diagnostic systems. *Comput. Chem. Eng.* **2000**, *24*, 785–791. [CrossRef]
796. Ganesan, R.; Das, T.; Venkataraman, V. Wavelet-based multiscale statistical process monitoring: A literature review. *IIE Trans.* **2004**, *36*, 787–806. [CrossRef]
797. Kadlec, P.; Grbić, R.; Gabrys, B. Review of adaptation mechanisms for data-driven soft sensors. *Comput. Chem. Eng.* **2011**, *35*, 1–24. [CrossRef]
798. Ding, S.; Zhang, P.; Jeinsch, T.; Ding, E.; Engel, P.; Gui, W. A survey of the application of basic data-driven and model-based methods in process monitoring and fault diagnosis. *IFAC Proc. Vol.* **2011**, *44*, 12380–12388. [CrossRef]
799. Sliskovic, D.; Grbic, R.; Hocenski, Z. Methods for Plant Data-Based Process Modeling in Soft-Sensor Development. *Automatika* **2011**, *52*, 306–318. [CrossRef]
800. Ge, Z.; Song, Z.; Gao, F. Review of Recent Research on Data-Based Process Monitoring. *Ind. Eng. Chem. Res.* **2013**, *52*, 3543–3562. [CrossRef]
801. Khatibisepehr, S.; Huang, B.; Khare, S. Design of inferential sensors in the process industry: A review of Bayesian methods. *J. Process Control* **2013**, *23*, 1575–1596. [CrossRef]
802. Yin, S.; Li, X.; Gao, H.; Kaynak, O. Data-Based Techniques Focused on Modern Industry: An Overview. *IEEE Trans. Ind. Electron.* **2015**, *62*, 657–667. [CrossRef]

803. Tidriri, K.; Chatti, N.; Verron, S.; Tiplica, T. Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. *Annu. Rev. Control* **2016**, *42*, 63–81. [CrossRef]
804. Shu, Y.; Ming, L.; Cheng, F.; Zhang, Z.; Zhao, J. Abnormal situation management: Challenges and opportunities in the big data era. *Comput. Chem. Eng.* **2016**, *91*, 104–113. [CrossRef]
805. Souza, F.A.A.; Araújo, R.; Mendes, J. Review of soft sensor methods for regression applications. *Chemom. Intell. Lab. Syst.* **2016**, *152*, 69–79. [CrossRef]
806. Reis, M.; Gins, G. Industrial Process Monitoring in the Big Data/Industry 4.0 Era: From Detection, to Diagnosis, to Prognosis. *Processes* **2017**, *5*, 35. [CrossRef]
807. Alauddin, M.; Khan, F.; Imtiaz, S.; Ahmed, S. A Bibliometric Review and Analysis of Data-Driven Fault Detection and Diagnosis Methods for Process Systems. *Ind. Eng. Chem. Res.* **2018**, *57*, 10719–10735. [CrossRef]
808. Wang, Y.; Si, Y.; Huang, B.; Lou, Z. Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008–2017. *Can. J. Chem. Eng.* **2018**, *96*, 2073–2085. [CrossRef]
809. He, Q.P.; Wang, J. Statistical process monitoring as a big data analytics tool for smart manufacturing. *J. Process Control* **2018**, *67*, 35–43. [CrossRef]
810. Gopaluni, R.B.; Tulsyan, A.; Chachuat, B.; Huang, B.; Lee, J.M.; Amjad, F.; Damarla, S.K.; Kim, J.W.; Lawrence, N.P. Modern Machine Learning Tools for Monitoring and Control of Industrial Processes: A Survey. *IFAC-PapersOnLine* **2020**, *53*, 218–229. [CrossRef]
811. Cohen, A.; Atoui, M.A. On wavelet based statistical process monitoring. *Trans. Inst. Meas. Control* **2020**, *44*, 014233122093570. [CrossRef]
812. Ahmed, Q.; Raza, S.A.; AlAnazi, D.M. Reliability based fault analysis models with industrial applications: A systematic literature review. *Qual. Reliab. Eng. Int.* **2020**, *37*, 1307–1333. [CrossRef]
813. Liu, Y.; Xie, M. Rebooting data-driven soft-sensors in process industries: A review of kernel methods. *J. Process Control* **2020**, *89*, 58–73. [CrossRef]
814. Curreri, F.; Fiumara, G.; Xibilia, M.G. Input Selection Methods for Soft Sensor Design: A Survey. *Future Internet* **2020**, *12*, 97. [CrossRef]
815. Jiao, Z.; Hu, P.; Xu, H.; Wang, Q. Machine Learning and Deep Learning in Chemical Health and Safety: A Systematic Review of Techniques and Applications. *ACS Chem. Health Saf.* **2020**, *27*, 316–334. [CrossRef]
816. Arunthavanathan, R.; Khan, F.; Ahmed, S.; Imtiaz, S. An analysis of process fault diagnosis methods from safety perspectives. *Comput. Chem. Eng.* **2021**, *145*, 107197. [CrossRef]
817. Taqvi, S.A.A.; Zabiri, H.; Tufa, L.D.; Uddin, F.; Fatima, S.A.; Maulud, A.S. A Review on Data Driven Learning Approaches for Fault Detection and Diagnosis in Chemical Processes. *ChemBioEng Rev.* **2021**, *8*, 239–259. [CrossRef]
818. Sun, Q.; Ge, Z. A Survey on Deep Learning for Data-Driven Soft Sensors. *IEEE Trans. Ind. Inform.* **2021**, *17*, 5853–5866. [CrossRef]
819. Wang, Y.; Qin, Y.; Lou, Z.; Ma, X. Survey on recursive statistical process monitoring methods. *Can. J. Chem. Eng.* **2022**, *100*, 2093–2103. [CrossRef]
820. Webert, H.; Döß, T.; Kaupp, L.; Simons, S. Fault Handling in Industry 4.0: Definition, Process and Applications. *Sensors* **2022**, *22*, 2205. [CrossRef]
821. Liu, S.; Lei, F.; Zhao, D.; Liu, Q. Abnormal Situation Management in Chemical Processes: Recent Research Progress and Future Prospects. *Processes* **2023**, *11*, 1608. [CrossRef]
822. Ramírez-Sanz, J.M.; Maestro-Prieto, J.A.; Arnaiz-González, A.; Bustillo, A. Semi-supervised learning for industrial fault detection and diagnosis: A systemic review. *ISA Trans.* **2023**, *143*, 255–270. [CrossRef]
823. Lou, C.; Atoui, M.A.; Li, X. Recent deep learning models for diagnosis and health monitoring: A review of research works and future challenges. *Trans. Inst. Meas. Control* **2023**, 014233122311571. [CrossRef]
824. Tien, D.X.; Lim, K.W.; Jun, L. Comparative study of PCA approaches in process monitoring and fault detection. In Proceedings of the 30th Annual Conference of IEEE Industrial Electronics Society, 2004. IECON 2004, Busan, Pepublic of Korea, 2–6 November 2004; Volume 3, pp. 2594–2599. [CrossRef]
825. Yin, S.; Ding, S.X.; Haghani, A.; Hao, H.; Zhang, P. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *J. Process Control* **2012**, *22*, 1567–1581. [CrossRef]
826. Zhang, K.; Hao, H.; Chen, Z.; Ding, S.X.; Peng, K. A comparison and evaluation of key performance indicator-based multivariate statistics process monitoring approaches. *J. Process Control* **2015**, *33*, 112–126. [CrossRef]
827. Jing, C.; Hou, J. SVM and PCA based fault classification approaches for complicated industrial process. *Neurocomputing* **2015**, *167*, 636–642. [CrossRef]
828. Askarian, M.; Escudero, G.; Graells, M.; Zarghami, R.; Jalali-Farahani, F.; Mostoufi, N. Fault diagnosis of chemical processes with incomplete observations: A comparative study. *Comput. Chem. Eng.* **2016**, *84*, 104–116. [CrossRef]
829. Rato, T.; Reis, M.; Schmitt, E.; Hubert, M.; De Ketelaere, B. A systematic comparison of PCA-based Statistical Process Monitoring methods for high-dimensional, time-dependent Processes. *AIChE J.* **2016**, *62*, 1478–1493. [CrossRef]
830. Zhang, K.; Ding, S.X.; Shardt, Y.A.; Chen, Z.; Peng, K. Assessment of T2- and Q-statistics for detecting additive and multiplicative faults in multivariate statistical process monitoring. *J. Frankl. Inst.* **2017**, *354*, 668–688. [CrossRef]
831. Fernandes, N.C.P.; Rato, T.J.; Reis, M.S. Modeling in the observable or latent space? A comparison of dynamic latent variable based monitoring methods for sensor fault detection. *Chemom. Intell. Lab. Syst.* **2022**, *231*, 104684. [CrossRef]

832. Zheng, J.; Zhao, C.; Gao, F. Retrospective comparison of several typical linear dynamic latent variable models for industrial process monitoring. *Comput. Chem. Eng.* **2022**, *157*, 107587. [CrossRef]
833. Hansen, H.H.; Kulachi, M.; Nielsen, B.F. Statistical process control versus deep learning for power plant condition monitoring. *Comput. Chem. Eng.* **2023**, *178*, 108391. [CrossRef]
834. Himmelblau, D. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*; Chemical Engineering Monographs; Elsevier: Amsterdam, The Netherlands, 1978.
835. Wang, X.Z. *Data Mining and Knowledge Discovery for Process Monitoring and Control*; Advances in Industrial Control; Springer: London, UK, 1999. [CrossRef]
836. Mason, R.L.; Young, J.C. *Multivariate Statistical Process Control with Industrial Applications*; SIAM: Philadelphia, PA, USA, 2002.
837. Ge, Z.; Song, Z. *Multivariate Statistical Process Control*; Advances in Industrial Control; Springer: London, UK, 2013. [CrossRef]
838. Ding, S. *Advanced Methods for Fault Diagnosis and Fault Tolerant Control*; Springer: Berlin/Heidelberg, Germany, 2021. [CrossRef]
839. Wang, Z.X.; He, Q.P.; Wang, J. Comparison of variable selection methods for PLS-based soft sensor modeling. *J. Process Control* **2015**, *26*, 56–72. [CrossRef]
840. Nespeca, M.G.; Pavini, W.D.; de Oliveira, J.E. Multivariate filters combined with interval partial least square method: A strategy for optimizing PLS models developed with near infrared data of multicomponent solutions. *Vib. Spectrosc.* **2019**, *102*, 97–102. [CrossRef]
841. Barra, I.; Mansouri, M.A.; Cherrah, Y.; Kharbach, M.; Bouklouze, A. FTIR fingerprints associated with a PLS-DA model for rapid detection of smuggled non-compliant diesel marketed in Morocco. *Vib. Spectrosc.* **2019**, *101*, 40–45. [CrossRef]
842. Sun, L.; You, G.; Cao, X.; Wang, M.; Ren, X. Comparative investigation for raw and processed Aconiti Lateralis Radix using chemical UPLC-MS profiling and multivariate classification techniques. *J. Food Drug Anal.* **2019**, *27*, 365–372. [CrossRef] [PubMed]