IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 3 September 2023, accepted 15 September 2023, date of publication 20 September 2023, date of current version 27 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3317516

RESEARCH ARTICLE

Advanced Statistical and Meta-Heuristic Based Optimization Fault Diagnosis Techniques in Complex Industrial Processes: A Comparative Analysis

FAIZAN E. MUSTAFA[1], ABDUL QAYYUM KHAN[1], (Senior Member, IEEE), ABDUS SAMEE[1], IJAZ AHMED[1], MUHAMMAD ABID[1], (Member, IEEE), MOHAMMAD M. ALQAHTANI[2], AND MUHAMMAD KHALID[3,4,5], (Senior Member, IEEE)

[1]Pakistan Institute of Engineering and Applied Sciences, Islamabad 44650, Pakistan
[2]Department of Industrial Engineering, King Khalid University, Abha 62529, Saudi Arabia
[3]Electrical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia
[4]IRC for Renewable Energy and Power Systems (IRC-REPS), King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia
[5]SDAIA-KFUPM Joint Research Center for Artificial Intelligence, Dhahran 31261, Saudi Arabia

Corresponding authors: Muhammad Khalid (mkhalid@kfupm.edu.sa) and Faizan E. Mustafa (faizanemustafa449@gmail.com)

Muhammad Khalid would like to acknowledge the support from KFUPM. Mohammad M. Alqahtani would like to acknowledge the support by the Deanship of Scientific Research through King Khalid University funded by the Large Group Research Project RGP2/239/44. The authors from PIEAS acknowledge the support from the Information Technology Endowment Fund.

ABSTRACT Industrial processes are nonlinear and complicated in nature, requiring accurate fault detection to minimize the deterioration in performance and to respond quickly to emergencies. This work investigates industrial process defect identification and isolation, which is analytically difficult owing to their complexity. This paper carefully analyzes four design methods for flaw identification and isolation based on Principal Component Analysis (PCA), Fisher Discriminant Analysis (FDA), Kernel Fisher Discriminant Analysis (KFDA), and Sequential quadratic programming (SQP). Our study includes the Tennessee Eastman Process (TEP) and the Penicillin Fermentation Process (PFP), among other comparable methods. We assess the proposed fault detection and isolation methods through detailed analysis and comparison. The simulation findings from our extensive investigation provide remarkable insights. Simulation findings show that FDA and KFDA work well in fault identification and isolation, but PCA has certain limits. We also considered SQP as a TEP fault detection and isolation improvement tool. SQP is noted for its success in nonlinear and restricted optimization problems, making it ideal for fault identification and isolation in complicated industrial processes. Data-driven design approaches increase problem identification in complicated industrial processes with greater reliability and efficiency than PCA-based methods. This study also shows that advanced data-driven techniques can improve industrial fault diagnosis, improving operational safety and system performance by leveraging the FDA, KFDA, and SQP.

INDEX TERMS Principal component analysis, fisher discriminant analysis, kernal fisher discriminant analysis, fault detection and isolation and meta-heuristic optimization.

I. INTRODUCTION

A. RESEARCH OBJECTIVE

Various engineering systems, like aeroplane engines, chemical processes, electronic equipment [1], and electrical machines [2], [3], [4], [5], need to be closely observed to make sure they work well, reliably, and safely [6]. Fault Detection and Diagnosis (FDD) is one of the most crucial technologies for assuring the safe operation of process industries in order to minimize performance degradation and avert potentially hazardous situations [7]. Although model plants are typically unavailable in the process industry, data-driven

The associate editor coordinating the review of this manuscript and approving it for publication was Shih-Wei Lin.

problems for fault detection and isolation strategies are a major topic of research [8].

The objective of this research is to assess the superiority of advanced data-driven techniques, exemplified by statistical, and meta-heuristic, over traditional-based methods for industrial fault diagnosis, emphasizing their ability to enhance problem identification, increase reliability, and achieve higher efficiency in addressing intricate industrial process issues, thereby advancing operational safety and performance.

B. LITERATURE REVIEW

A fault can be characterized as a divergence from the normal distinguishing property that is not permitted [9]. For example, a number of different faults, such as equipment degradation, seasonal changes, plant-wide oscillations, and several others, can cause a variety of malfunctions in the process equipment. In most cases, the actuators, components, and sensors are the ones to go wrong. These defects are responsible for a variety of industrial incidents that have occurred in the sector as well as emergency shutdown scenarios [10]. Researchers around the globe have employed a wide variety of strategies to diagnose problems for process breakdowns and malfunctions in pursuit of an efficient Fault Detection and Diagnosis (FDD) [11]. The term "Advanced Statistical and Meta-heuristic methods" encompasses a collection of complex techniques employed in the FDD domains for data analysis, optimization, and problem-solving. The specificity's of these strategies are outlined below:

1) Advanced Statistical Methods: These approaches transcend the scope of fundamental statistical analysis. These approaches encompass increasingly intricate and specialized techniques for investigating, examining, and deriving meaningful conclusions from data. The aforementioned techniques frequently address the analysis of extensive datasets, multidimensional domains, and complex, non-linear associations. Advanced statistical approaches encompass a range of sophisticated techniques that are employed to analyze complex data sets and derive meaningful insights. Some illustrative examples of such methods include:
   - Machine Learning Algorithms: Algorithms such as neural networks, random forests, support vector machines, and gradient-enhancing are commonly employed in many computational tasks, including classification, regression, clustering, and others.
   - Bayesian Inference: The application of a probabilistic framework to the process of statistical inference, wherein assumptions are iteratively revised and updated in response to the acquisition of new evidence [12].
   - Dimensionality Reduction: PCA and t-SNE are widely used techniques in data analysis that facilitate the reduction of variables while preserving crucial information.

2) Meta-heuristic: The optimization algorithms discussed below draw inspiration from natural processes, social behavior, and mathematical models. Complex optimization issues may be effectively addressed with these methods, since they are capable of overcoming the limitations and time constraints associated with traditional approaches. Meta-heuristic techniques are specifically developed to effectively navigate through an extensive search space in order to identify solutions that are close to optimum. Meta-heuristic approaches encompass a variety of techniques that are employed to solve complex optimization problems [13], [14], [15], [16]. Some notable examples of these methods include:
   - Genetic Algorithms: Algorithms that employ the principles of natural selection to iteratively generate and refine solutions for optimization purposes.
   - Tabu Search: One approach to optimization involves the utilization of a short-term memory mechanism that keeps track of previously visited solutions, so preventing redundant revisits and promoting a more varied exploration of the search space.
   - Simulated Annealing: An optimization strategy that draws inspiration from the annealing process in metallurgy, when a material undergoes a gradual cooling procedure to mitigate the presence of flaws. The process entails systematically and repeatedly examining the range of potential solutions, even if they are not the most optimum, in order to avoid being trapped in local optima.

Advanced statistical and meta-heuristic techniques are of paramount importance in addressing intricate challenges across diverse fields such as engineering, research, business, and computer science. These methodologies facilitate the management of extensive data analysis, optimization, and decision-making obstacles that conventional techniques may find difficult to tackle.

In the industrial domain, interconnections between processes are prevalent, and system issues often trigger a multitude of alarms simultaneously [17], [18]. Consequently, the use of univariate statistical monitoring techniques becomes inadequate for higher-dimensional systems. The presence of numerous alarms fails to offer a clear indication of the variables responsible for the anomaly. In contrast, multivariate statistical methods prove to be effective in fault detection as they account for the correlations between variables [19].

Several techniques, including data-driven minimization techniques, have been explored to address the challenges posed by faults in industrial systems [20], [21]. Among these strategies, Principal Component Analysis (PCA) emerges as a prominent approach. PCA is renowned for its capability to reduce the dimensionality of data while preserving crucial information from the original dataset. Over the course of more than four decades, PCA has garnered extensive attention and utilization in the industrial domain due to its profound dimensionality reduction capabilities [22]. PCA serves as a valuable tool in transforming high-dimensional data into a lower-dimensional representation while preserving vital information [23]. In the realm of fault diagnosis, Hotelling's $T^2$and Square Prediction Error (SPE) Q statistics are frequently employed in conjunction with PCA. The T^2 statistics assess the PCA subspace, providing insights into deviations from the normal behavior, while the Q statistics offer an evaluation of the residual space, capturing anomalous patterns. PCA provides useful insights into the structures of data and assists in simplifying complicated datasets. However, it is crucial to realize the inherent constraints associated with PCA. The aforementioned constraints are especially evident when examining the attributes of unprocessed data, since they can have a substantial impact on the efficacy of PCA. A major shortcoming of PCA is the linearity assumption. PCA assumes that data linkages are linear, therefore linear combinations of the original characteristics may properly reflect data changes. In real life, data often have non-linear correlations that PCA cannot represent. PCA may not accurately describe complicated non-linear patterns, resulting in inadequate variance explanation and dimensionality reduction. Another limitations of PCA is outlier sensitivity. Outliers can disproportionately affect PCA principle components. Additionally, outliers can affect PC, misrepresenting the data's structure as PCA maximizes variance.

Fisher Discriminant Analysis (FDA) emerges as a valuable technique for nonlinear dimensionality reduction. Its core functionality involves the calculation of a transformation matrix that selects a set of vectors aimed at maximizing inter-class separation while minimizing intra-class separation [24], [25]. Notably, FDA excels in fault isolation diagnostics, showcasing its effectiveness in identifying and isolating specific faults within a system. Conversely, PCA demonstrates its prowess in fault detection [26], [27], [28].

Kernel Fisher Discriminant Analysis (KFDA) serves as a remarkable technique for dimensionality reduction, specifically tailored for highly nonlinear systems. Unlike FDA, KFDA exhibits significantly enhanced fault isolation capabilities in the context of nonlinear systems. KFDA accomplishes dimensionality reduction by estimating a dataset's projection into a higher-dimensional space using a linear classification algorithm within its feature space. However, the computational complexity associated with the implicit feature space can be laborious. To overcome this challenge, the kernel method is employed, circumventing the computational burden while maintaining the efficacy of KFDA [29].

Metaheuristic optimization is a highly effective approach that employs intelligent algorithms to efficiently address intricate optimization problems, surpassing conventional methods [30], [31]. The applications of this methodology are wide-ranging, encompassing diverse fields such as engineering, finance, logistics, and data science [29], [32], [33], [34], [35]. It facilitates effective resource allocation, parameter tuning, and decision-making procedures. Among the plethora of optimization approaches, Sequential Quadratic Programming (SQP) emerges as a formidable contender, renowned for its prowess in industrial process optimization owing to its diminished computational time, seamless implementation, and remarkable efficacy in tackling intricate constraint problems [36], [37]. Various shortcomings in different systems can be identified and diagnosed with the use of optimization techniques, which have found widespread use in this area. Faults in the transmission network, for instance, can be detected and localized using optimization-based fault detection algorithms [38] in the power systems industry. Optimization methods have also been used in the chemical industry to enhance both the effectiveness and dependability of complex system problem detection and diagnosis [39]. These implementations show how optimization strategies can improve system speed and security through fault detection.

Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) are commonly employed by researchers for the purposes of dimensionality reduction and classification. However, the applicability of these methods in complicated industrial processes is sometimes hindered by certain restrictions. The LDA algorithm assumes a Gaussian distribution and equal class covariance, which might limit its effectiveness when applied to heterogeneous and imbalanced data commonly seen in industrial environments [40]. Likewise, the performance of SVMs may be adversely affected by the curse of dimensionality and the difficulties associated with accurately defining complex decision boundaries that are inherent in sophisticated datasets commonly seen in industrial settings [41]. On the other hand, the selected methodologies, namely PCA, FDA, KFDA, and SQP, have been specifically designed to effectively tackle these intricacies. PCA is highly effective in capturing the fundamental variance present in a dataset. However, it may potentially disregard significant variances that are peculiar to certain classes. The FDA algorithm improves the distinguishability between different classes, however it may encounter difficulties in accurately capturing nonlinear associations. The integration of KFDA and SQP combines the benefits of PCA, FDA, and nonlinear classification in order to overcome the limits associated with fault identification in complex industrial processes, resulting in improved accuracy.

C. CONTRIBUTIONS

Within the context of this review, the main aim is to provide a thorough and complete assessment of the latest developments in defect detection algorithms specifically pertaining to industrial processes. The primary objective of this research activity has been to devise efficacious methodologies for the identification and containment of faults, particularly in situations when the lack of a dependable mathematical model presents difficulties. The review's contributions may be succinctly described as follows:

1) Comprehensive Framework Development: The primary focus of the present research is the formulation and establishment of an all-encompassing framework

specifically designed for the purpose of detecting and isolating faults in industrial processes. The proposed framework not only tackles the challenges that arise due to the lack of a mathematical model but also places significant emphasis on the implementation of data-driven design techniques, hence enhancing its versatility across many circumstances.

2) Integration of Advanced Techniques: In this study, we reviewed several sophisticated methodologies, such as PCA, FDA, KFDA, and SQP, inside the established framework. The selection of these approaches was based on their well-established relevance in the field of defect identification.
3) Comparative Analysis of Fault Isolation: An essential component of this review article is a comprehensive comparative examination of the fault isolation performance across the employed methodologies. It is worth noting that although PCA had considerable ability in detecting faults, its effectiveness in isolating faults was restricted when compared to the superior performances of FDA, KFDA and SQP.

Throughout this research endeavor, a comprehensive framework for detecting and isolating faults in industrial processes has been developed, considering the absence of a mathematical model and necessitating the application of data-driven design strategies. Notably, PCA, FDA, KFDA, and SQP have been implemented as part of this framework. To evaluate the effectiveness of these methods, renowned benchmark processes such as the Tennessee Eastman Process (TEP) and the Penicillin Fermentation Process (PFP) have been meticulously examined. The findings of this study reveal that each method exhibits the capability to detect faults with relative ease. However, PCA falls short in terms of fault isolation when compared to FDA and KFDA. This disparity in performance can be attributed to PCA's limited consideration of information from other classes, whereas both FDA and KFDA take such information into account. Consequently, FDA and KFDA demonstrate significantly superior fault isolation capabilities compared to PCA, aligning with the observed results and highlighting the importance of considering information from multiple classes. In contrast, SQP has emerged as a notably proficient methodology, surpassing approaches such as KFDA and PCA in terms of fault detection accuracy and false alarm mitigation. Moreover, the convergence behavior of SQP exhibits an exceptional characteristic, enabling it to rapidly attain convergence to the desired solution, thereby minimizing computational time and enhancing the overall efficiency of fault detection algorithms.

These findings contribute to the growing understanding of fault detection and isolation in industrial processes, shedding light on the advantages and limitations of different data-driven design strategies. The observed performance disparities underscore the significance of utilizing advanced techniques, such as FDA, KFDA, and SQP which consider the correlation and inter-class information, leading to improved fault isolation outcomes. By leveraging these insights, researchers and practitioners can enhance the reliability and effectiveness of fault detection and isolation, ultimately bolstering the operational safety and performance of industrial systems.

D. CHARACTERISTICS AND DIFFICULTIES OF METHODS

In this article, PCA, FDA, KFDA and SQP approaches are considered which are designed for fault diagnosis of industrial processes. In particular, these approaches are applied to two benchmark processes for performance comparison. The main characteristics and difficulties for fault diagnosis are as follows:

1) **Fault detection** The ability to diagnose a faulty condition correctly and quickly is crucial and highly desired. The main difficulty in fault detection is false alarms which should be minimal. In this article, PCA is used for fault detection which is a reliable, efficient, and accurate algorithm.
2) **Fault isolability** The ability of the diagnostic system to distinguish between various failures is known as isolability. The main hurdle in achieving fault isolation is that transformation matrix orientation is achieved by hit and trial technique which should be observed carefully. In this research, FDA and KFDA are well suited for the isolation of faults.
3) **Robustness** A fault diagnosis system should be robust and be able to withstand uncertainties and disturbances which is done by maintaining a False Alarm Rate (FAR) less than 2% in this article. The main difficulty in achieving a lower false alarm rate is that the system should be sensitive to faults in such a way that false alarms should be minimal.
4) **Fault identification** Another issue in fault diagnosis is that finding the major root cause of the fault is necessary when undergoing thorough fault diagnosis. SQP plays a major role in fault identification by optimization of the TEP process through metaheuristic approaches.
5) **Sensitivity** One of the main issues of fault diagnosis system is that it should be sensitive to faults. It is observed by simulation by case studies of TEP and PFP that they are sensitive to faults as they are detected when introduced in a process. The main hurdle in making a system sensitive is that the threshold level should be adaptive and robust.

E. ORGANIZATION

The structure of the paper is organized as follows. Section II provides a concise overview of Principal Component Analysis (PCA). In Section III, Fisher Discriminant Analysis (FDA) is comprehensively discussed, highlighting its underlying principles and methodologies. The detailed exploration of Kernel Fisher Discriminant Analysis (KFDA) is presented in Section IV, providing a deeper understanding of this advanced technique. The modelling of SQP for fault detection and false alarm minimization is discussed in Section V. Furthermore, Section VI showcases the practical application of the four strategies on various real-world systems, illustrating their effectiveness and performance in fault detection and isolation scenarios. The case studies shed light on the practical implications and advantages of employing PCA, FDA, KFDA, and SQP in industrial settings. Finally, Section VII summarizes the key findings and contributions of the paper, emphasizing the significance of data-driven design strategies in fault detection and isolation for industrial applications. In this manner, the present investigation provides a comprehensive examination of PCA, FDA, KFDA, and SQP presents case studies highlighting their practical deployment, and concludes with a concise summary of the research outcomes.

F. ACRONYMS, UNITS AND DIMENSIONS

NOMENCLATURE

PCA	Principal Component Analysis.
FDA	Fisher Discriminant Analysis.
KFDA	Kernal Fisher Discriminant Analysis.
SPE	Sqaure Prediction Error.
SQP	Sequential Quadratic Programming.
Z_{obs}	Input-output data set.
$\bar{Z}_{obs}$	Mean of input-output data.
Z(k)	Normalized data set $\in (m \times n)$.
Ψ	Number of samples.
Ψ	Number of principal components.
C	Covariance of matrix $\in (m \times m)$.
P	Loading vector $\in (m \times m)$.
Γ	Diagonal matrix of covariance matrix eigen values $\in (m \times m)$.
m	Number of input-output variables.
Γ_{pc}	Principal component of eigen values $\in (m \times \psi)$.
Γ_{res}	Residual component of eigen values $\in (m \times m - \psi)$.
P_{pc}	Principal component of loading vector $\in (m \times \psi)$.
P_{res}	Residual component of loading vector $\in (m \times m - \psi)$.
T	Transformation matrix.
$J_{th,SPE}$	Threshold for square prediction error.
$J_{th,T^2_{PCA}}$	Threshold for T^2 statistics.
T^2	T^2 statistics.
S_t	Total scattering matrix $\in (m \times m)$.
X	Input-output data set $\in (n \times m)$.
x_i	Transpose of ith row of input-output data set $\in (m \times n)$.
$\bar{x}$	Mean of input-output data set $\in (m \times n)$.
S_w	Within scattering matrix $\in (m \times m)$.
S_j	Scattering matrix of class j $\in (m \times m)$.
S_b	between Scattering matrix $\in (m \times m)$.
λ	Eigen vector for generalized vector solution.
W_ψ	Set of eigen vectors.
w_k	Matrix of eigen vectors.
T^2_δ	T^2_δ threshold statistics.
ζ	Number of classes.
Φ	Non linear function for initial data.
V_i	Matrix of eigenvectors for KFDA approach.
Z_i	Transformation matrix for KFDA approach.
α	Combination of vectors for KFDA approach.

II. PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA stands as a powerful multivariate statistical technique that facilitates the transformation of correlated variables within the original data into a set of uncorrelated variables, encapsulating the most salient information [42]. Its roots trace back to pioneering work by Pearson [43], while the modern conceptualization of PCA was solidified by Hotelling [44], who introduced the notion of Principal Components (PCs). The amalgamation of these historical and contemporary contributions has led to PCA's prominent position in data analysis and dimensionality reduction methodologies. PCA endeavors to establish a novel coordinate system that emphasizes the salient dimensions, distilling essential information from the original data [45]. By employing a linear transformation, PCA effectively partitions the initial dataset into two subspaces: the major subspace capturing the most significant variations, and the residual subspace containing the remaining information [46], [47]. This unique technique excels in reducing the dimensionality of data, thus enabling efficient data representation. The selection of the critical number of PCs plays a pivotal role in capturing the dominant features of the dataset.

The PCs serve multiple essential objectives, including:

1) **Extraction of Crucial Information:** PCs aim to capture the most vital information embedded within the data table. By identifying the dominant patterns and variations, they distill the fundamental insights that drive the underlying phenomena.
2) **Dimensionality Reduction:** An intrinsic goal of PCs is to minimize the size of the dataset while retaining the critical information necessary for analysis. By eliminating redundant or less informative dimensions, PCs facilitate efficient data representation, leading to enhanced computational efficiency and interpretability.
3) **Simplification of Data Description:** PCs offer a means to simplify the representation and description of complex datasets. By condensing the information into a reduced set of dimensions, PCs provide a concise summary that encapsulates the essential characteristics of the data, enabling concise and insightful data exploration.
4) **Structural Analysis of Samples and Variables:** PCs enable a comprehensive examination of the interrelationships between samples and variables. Through their calculation, PCs unveil the underlying structure and dependencies within the dataset, facilitating the identification of key factors that contribute to the observed patterns and variations.

By pursuing these objectives, PCs serve as a powerful tool for data analysis, enabling researchers to uncover crucial information, reduce dataset size, simplify data representation, and delve into the structural relationships within the dataset.

PCA effectively addresses these objectives by computing new variables, referred to as principal components. These components are generated as linear combinations of the original variables, aimed at achieving specific goals. The primary component exhibits the largest variance, thereby capturing the most significant variation within the dataset. Moreover, PCA offers a visual representation of the data and variables in the form of maps, illustrating the patterns of similarity between observations and variables. This visualization enables researchers to gain insights into the underlying structure and relationships present in the dataset. The process of computing the PCs and subsequently transforming the data based on these components is known as PCA. Often, only a subset of the PCs is utilized, while the remaining components are disregarded. Numerous disciplines, ranging from population genetics, micro biome research, atmospheric science, to image processing, rely on PCA to unravel patterns and derive meaningful information from complex datasets. PCA serves as a fundamental analytical tool in these domains, facilitating data exploration and dimensionality reduction [48]. The basic structure of the PCA being discussed is as follows:

- Step 1: To initiate the analysis, consider a measurement vector of dimensionality m and a total of Ψ samples for a given process, as depicted in Equation 1 [49]:

$$\begin{bmatrix} Z_{obs,1} \\ , \\ .. \\ Z_{obs,m} \end{bmatrix} \in R^m \tag{1}$$

  In equation (1), Z_{obs} and Ψ represents the vector is an observation vector and samples, each associated with a unique measurement vector, respectively. This initial setup provides the foundation for subsequent data analysis and exploration, as various techniques and methodologies can be applied to uncover valuable insights from the dataset.
- Step 2: This subsequent step involves normalizing the vector by computing the mean and standard deviation of the observation vector, denoted as Z_{obs}. This normalization process is governed by the equations 2 and 3 [49] presented below:

$$\overline{Z}_{obs,i} = \frac{1}{\Psi}\sum_{j=1}^{\Psi} Z_{obs,i}(j) \tag{2}$$

$$\sigma^2_{obs,i} = \frac{1}{\Psi-1}\sum_{j=1}^{\Psi}(Z_{obs,i}(j) - \overline{Z}_{obs,i}(j))^2 \tag{3}$$

  In equation (2), $Z_{obs,i}$ represents the mean value of the jth component, calculated by summing all corresponding values across the Ψ samples and dividing by the total number of samples. Equation (3) defines $\sigma_{obs,i}$ as the standard deviation of the jth component, computed by subtracting the mean value of the jth component from each sample value, squaring the differences, summing them up, dividing by $(\Psi - 1)$, and taking the square root.
- Step 3: To determine the location of the normalized vector, we utilize the following relation:

$$Z(k) = \begin{bmatrix} \frac{Z_{obs,1}(k) - \overline{Z}_{obs,1}}{\sigma_{obs,1}}. \\ . \\ . \\ \frac{Z_{obs,m}(k) - \overline{Z}_{obs,m}}{\sigma_{obs,m}} \end{bmatrix} \tag{4}$$

  while $Z(k)$ is the normalized matrix with process variables and samples. Normalized vector can be summarized as,

$$Z(k) = \begin{bmatrix} Z(1) & . & . & , Z(\Psi) \end{bmatrix} \in R^{m\times\Psi} \tag{5}$$

  Here, Z represents the normalized vector, obtained by subtracting the mean vector from the observation vector (Z_{obs}) and dividing the result by the standard deviation vector (σ). This transformation ensures that the normalized vector is centered around zero mean and scaled to unit variance, facilitating meaningful and standardized comparisons within the dataset.
- Step 4: To take covariance of equation (5):

$$C = \frac{Z \times Z^T}{\Psi - 1} \tag{6}$$

  where C is the covariance of $Z \times Z^T$ matrix.
- Step 5: To apply Singular Value Decomposition (SVD) as shown below:

$$\frac{Z \times Z'}{\Psi - 1} = P\Gamma P^T \tag{7}$$

  where P is the loading vector and Γ is the singular matrix with eigen values.

$$\Gamma = diag\left(\gamma_1^2, \ldots \gamma_m^2\right), \quad \gamma_1^2 \geq \gamma_2^2 \geq \gamma_3^2 \geq \ldots \gamma_m^2 \tag{8}$$

  where γ_1 is the first eigen value.
- Step 6: To determine the number of PCs ψ and divide P and Γ into the following:

$$\Gamma = \begin{bmatrix} \Gamma_{pc} & 0 \\ 0 & \Gamma_{res} \end{bmatrix}, \quad \Gamma_{pc} = diag\left(\gamma_1^2, \ldots \gamma_\psi^2\right) \tag{9}$$

$$\Gamma_{res} = diag\left(\gamma_{\psi+1}^2, \ldots \gamma_m^2\right) \in R^{(m-\psi)\times(m-\psi)}, \quad \gamma_\psi^2 \geq\geq \gamma_{\psi+1}^2. \tag{10}$$

$$P = \begin{bmatrix} P_{pc} & P_{res} \end{bmatrix} \in R^{m\times m}, \; P_{pc} \in R^{m\times\psi}, \; P_{res} \in R^{\psi\times m} \tag{11}$$

  where P_{pc} is the principal component of loading vector while P_{res} is the residual component of loading vector. Similar is the case for Γ matrix.
- Step 7: Find the transformation matrix:

$$T = Z^T \times P \tag{12}$$

  where T is the transformation matrix.

The tracking indices used for Principal Component Analysis (PCA) include Hotelling's statistics, such as the Square Prediction Error (SPE) or Q statistics, defined by the equation below:

$$J_{th,SPE} = \theta_1\left(\frac{c_\alpha\sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2}\right)^{\frac{1}{h_0}} \tag{13}$$

$$\theta_i = \sum_{j=\psi+1}^{m}\left(\gamma_j^2\right)^i, i = 1,2,3, h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \tag{14}$$

c_β is the normal deviate corresponding to β percentile. where $J_{th,SPE}$ is the threshold for SPE. Similarly, $J_{th,T^2_{PCA}}$ statistic threshold is given by:

$$J_{th,T^2_{PCA}} = \frac{\psi\left(\Psi-1)(\Psi+1)\right)}{\Psi\left(\Psi-1\right)}F_\beta(\psi,\Psi-\psi) \tag{15}$$

where F_β is the F-distribution with ψ and $\Psi-\psi$ degrees of freedom.

T^2 and Q are calculated by the following equations,

$$SPE_{PCA} = \left\|\left(I - P_{pc}P_{pc}^T\right)z\right\|_E^2 \tag{16}$$

$$T^2_{PCA} = z^T P_{pc}\Gamma_{pc}^{-1}P_{pc}^T z \tag{17}$$

The Q statistic serves as a measure of the deviation or distance of a given observation from the normal behavior captured by the PCA model. By evaluating the magnitude of the Q statistic, deviations or abnormalities in the dataset can be detected, aiding in fault detection and identification. On the basis of following set of rules, fault is detected, $SPE \le J_{th,SPE_{PCA}}$ and $T^2_{PCA} \le J_{th,T^2_{PCA}} \Rightarrow$ fault-free, otherwise faulty.

III. FISHER DISCRIMINANT ANALYSIS (FDA)

While PCA is a renowned technique for dimensionality reduction, it does not explicitly consider the information pertaining to different types of faults during the development of its transformation matrix, denoted as T. However, this aspect is addressed by FDA. In FDA, the transformation matrix is constructed with the specific objective of minimizing within-class scatter while simultaneously enhancing the separation between different fault classes. By considering the interplay between classes, FDA aims to extract discriminant features that effectively distinguish and classify different types of faults. This discriminative power allows FDA to outperform PCA when it comes to fault detection and identification tasks, particularly in scenarios where the differentiation between different fault classes is crucial.

Let us consider a process characterized by m variables, with Ψ observations available for each variable. Furthermore, assume that there are ζ distinct fault classes affecting the process. For the jth class, let Ψ_j represent the number of observations, which are organized and stacked in a matrix denoted as $x \in \mathbb{R}^{\Psi\times m}$. In this matrix, each row corresponds to the transpose of the ith observation, denoted as x_i.

- Step 1: Calculate total-scatter matrix S_t [50]:

$$S_t = \sum_{i=1}^{\Psi}(x_i-\bar{x})(x_i-\bar{x})' \tag{18}$$

where x_i is the transpose of i^{th} row of matrix $x \in R^{\Psi * m}$, Ψ is the number of total samples and $\bar{x}$ means the mean of x.

- Step 2: Calculate within scattering matrix S_w:

$$S_w = \sum_{j=1}^{\zeta} S_j \tag{19}$$

$$S_j = \sum_{x_i\in x_j}^{\Psi_j}(x_i-\bar{x}_j)(x_i-\bar{x}_j)' \tag{20}$$

where S_j is the scattering matrix of class j,Γ_j is the number of samples of class j, x_i belong to x_j and $\bar{x}_j$ is the mean of x_j.

- Step 3: To determine the between class scattering matrix S_b [50],

$$S_b = \sum_{j=1}^{\zeta}\Psi_j(\bar{x}_j-\bar{x})(\bar{x}_j-\bar{x})' \tag{21}$$

$$S_t = S_b + S_w \tag{22}$$

where S_t is the total scattering matrix.

- Step 4: Find the FDA vectors by following relation:

$$S_b w_k = \lambda_k S_w w_k \tag{23}$$

where w_k is the matrix of eigen vector. The purpose of the primary FDA vector is to maximize the diffuse between classes whereas minimizing the scatter inside classes:

$$opt(\frac{w_k^T S_b w_k}{w_k^T S_w w_k}) \tag{24}$$

whereas it is assumed that the inverse of S_w exists.

- Step 5: Separate the eigenvectors:

$$W_\psi = \begin{bmatrix} w_1 & w_2 & w_3 & . & .w_\psi \end{bmatrix} \tag{25}$$

- Step 6: FDA transformation vectors are calculated by [50]:

$$Z_i = W_\psi^T \times x_i \tag{26}$$

where Z_i is the transformation matrix.

For fault detection of any class of faults T^2 statistics is used [50],

$$T_j^2 = x^T \times W_{\bar{\psi}}(W_{\bar{\psi}}^T S_j W_{\bar{\psi}})^{-1} W_{\bar{\psi}^T} \times x \tag{27}$$

Threshold for j_{th} fault is computed as follows [50],

$$T_\delta^2 = \frac{\bar{\psi}(\Psi-1)(\Psi+1)}{\Psi(\Psi-1)}F_\beta(\bar{\psi},\Psi-\bar{\psi}) \tag{28}$$

If $T_j^2 > T_\delta^2$ then there is fault, otherwise no fault.

IV. KERNAL FISHER DISCRIMINANT ANALYSIS (KFDA)

In statistical analysis, a nonlinear extension of linear discriminant analysis is known as Kernel Fisher Discriminant Analysis (KFDA). KFDA aims to minimize within-class scatter while maximizing between-class scatter, thereby enhancing the discriminative power of the analysis. To overcome the limitations posed by the implicit feature space, a nonlinear variant of FDA called KFDA has been proposed in the literature [51].

KFDA employs a simple nonlinear mapping to project a dataset exhibiting nonlinear behavior onto a higher-dimensional feature space. Subsequently, a linear classification method is applied within this feature space. However, due to the potentially enormous or infinite dimensions of the implicit feature space, direct computations in this space are impractical. To address this challenge, KFDA utilizes a kernel function to calculate the dot products of vectors in the feature space, rather than operating directly in the implicit feature space. The KFDA algorithm typically involves three primary steps. Firstly, the KFDA vectors are computed through offline computations using input-output data. Subsequently, KFDA transformation vectors are derived, allowing for the projection of data onto the transformed feature space. Finally, data is classified based on the derived KFDA vectors, employing the discriminative power of the method [52]. This approach enables effective nonlinear analysis and classification, addressing the limitations associated with traditional linear discriminant analysis techniques.

- Step 1: The KFDA vectors are obtained by leveraging the training data, which is partitioned into distinct classes [53]. This division allows for the identification of specific groups or categories within the dataset, each representing a distinct class:

$$x = \begin{bmatrix} x_1 & x_2 & . & . & .x_\zeta \end{bmatrix} \tag{29}$$

  where ζ is the number of classes. Through the use of non-linear function, Φ the initial data is converted into high dimensional space,

$$\Phi : x_i \to \Phi(x_i) \in F^h, \ h > m$$
$$V_i = \sum_{i=1}^{\Psi} (\alpha_i \Phi(x_i)) \tag{30}$$

  whereas V_i is the matrix of eigenvectors, α_i is the combination of vectors $\alpha_1, \alpha_2, \ldots \alpha_\Psi$.
- Step 2: The KFDA transformation vectors are subsequently calculated to enable the projection of online measurements into a higher-dimensional space. This transformation process plays a crucial role in extending the analysis beyond the original measurement space, facilitating the exploration of complex relationships and capturing intricate patterns. By utilizing the KFDA transformation vectors, the online measurements are effectively mapped into the transformed feature space. This higher-dimensional representation allows for a more comprehensive analysis, enabling the detection of subtle variations and discriminating characteristics that may not be apparent in the original measurement space. Through this transformation, the KFDA algorithm enhances the ability to capture and leverage the discriminative information present in the online measurements, opening up new avenues for accurate classification, fault detection, and performance analysis [53]:

$$Z_i = V_i \Phi(x_i) \tag{31}$$

- Step 3: The final step in KFDA involves the classification of the transformation matrix [54], [55]. After the online measurements are projected into the higher-dimensional feature space using the KFDA transformation vectors, the classification process takes place.

V. SEQUENTIAL QUADRATIC PROGRAMMING (SQP)

The TEP and the PFP are two well-known industrial processes that make extensive use of SQP. Taking into account the inherent complexities of fault identification and diagnosis, SQP has shown to be an excellent optimization tool in the TEP. The complex restrictions of the TEP are easily managed by SQP since the problem is formulated as a restricted optimization job; this allows for precise fault detection and efficient process control. SQP has also shown great success in the PFP, where it has been used to optimize the parameters of the fermentation process, leading to increased penicillin output and better overall efficiency. Due to its iterative nature, SQP is able to efficiently explore the PFP's multi-dimensional search space and converge on the parameters that would produce the highest yield of penicillin. These use cases demonstrate how effective SQP is at meeting the challenges of real-world industrial processes and how it can be optimized and controlled in a wide variety of settings. Various defects in different systems can be identified and diagnosed with the use of optimization techniques, which have found widespread use in this area. Faults in transmission lines, for instance, can be detected and localized using optimization-based fault detection algorithms in the power systems industry. Optimization methods have also been used in the chemical industry to enhance the efficiency and reliability of complex system problem detection and diagnosis. These implementations show how optimization strategies can improve system speed and security through fault detection.

SQP effectively searches the TEP's multi-dimensional search space because of its iterative nature. SQP iteratively updates the variables to enhance the objective function and computes the search direction by approximating the Hessian matrix. Exploration and exploitation are balanced in SQP, allowing for effective fault identification by using first- and second-order derivative information. Particularly helpful for the TEP is SQP's capability to handle nonlinear constraints, which permits more precise modelling of the process's dynamics and restrictions. Constraints like material balance equations and equipment operating restrictions are taken into account while the process variables are optimized.

This all-encompassing method not only enhances the reliability of issue detection, but also aids in keeping the TEP running smoothly and effectively. In addition, SQP is preferable due to its convergence qualities, which aid in defect identification and alert minimization in the TEP. In a short amount of time, it converges on the optimal solution, allowing for immediate defect diagnosis and effective action. SQP is useful for fault identification and alarm minimization, improving the security and dependability of TEP operations thanks to its efficient handling of the TEP's complicated dynamics and its convergence characteristics.

In order to address the issue of minimizing false alarms and detecting faults using SQP, the following procedural steps are undertaken:

A. PROBLEM DESCRIPTION

- **Decision Variables:** Let $\mathcal{L}$ denote the vector of decision variables that will be used to optimise the system.
- **Constraints:** Constraints, such as physical boundaries, operational restrictions, and sensor boundaries of TEP, will be defined now to guarantee the system stays inside the predetermined parameters.
- **Objective Function:** The objective function is a mathematical representation that measures both the frequency of false alarms and the accuracy of the identification of faults. The expression $\mathbf{f}(\mathcal{L})$ is used to denote the objective function that is to be minimized.

B. MODELLING THE OBJECTIVE FUNCTION

The primary aim of the objective function should be to strike a delicate equilibrium between the minimization of false alarms and the maximization of fault detection accuracy. The formulation of this concept can be expressed as:

$$\mathbf{m}inf(\mathcal{L}) = \omega_1 \times \text{FAR}(\mathcal{L}) + \omega_2 \times \text{FDR}(\mathcal{L}). \tag{32}$$

In equation (32), False Alarm Rate (FAR)$\mathcal{L}$ represents the occurrence of false alarms, Fault Detection Rate (FDR)$\mathcal{L}$ represents the fault detection accuracy, and ω_1 and ω_2 are weighting factors that determine the relative importance of false alarms and fault detection. The current challenge at the moment involves the development of a system model that accurately encompasses the various dynamics and behaviours exhibited by the system. This model can be effectively represented through a collection of formulas, which serve to describe its complex workings:

$$g(\mathcal{L}) = 0, \quad h(\mathcal{L}) \leq 0. \tag{33}$$

C. LOGICAL STEPS OF THE SQP OPTIMIZATION ALGORITHM

- **Step 1:** The SQP algorithm employs an iterative approach for continuously updating the decision variables in order to minimize the objective function, while simultaneously ensuring that the limitations are satisfied.
- **Step 2:** During every cycle, SQP algorithm solves a subproblem of quadratic programming by employing the estimated values of the gradient $\nabla f(\mathcal{L})$ and Hessian matrix $(H(\mathcal{L}))$.
- **Step 3:** The subproblem can be formulated as:

$$\begin{aligned} &\min \Delta\mathcal{L}\left[\nabla f(\mathcal{L})^{\tau}\Delta\mathcal{L} + 0.5(\Delta\mathcal{L}^{\tau})H(\mathcal{L})\,\Delta\mathcal{L}\right],\\ &\text{subject to } g(\mathcal{L}) + Jg(\mathcal{L})\,\Delta\mathcal{L} = 0,\\ &\qquad h(\mathcal{L}) + Jh(\mathcal{L})\,\Delta\mathcal{L} \leq 0. \end{aligned} \tag{34}$$

  In equation (34), symbol $\Delta\mathcal{L}$ denotes the variation in decision variables, while $Jg(\mathcal{L})$ and $Jh(\mathcal{L})$ refer to the Jacobian matrices of $g(\mathcal{L})$ and $h(\mathcal{L})$ correspondingly.
- **Step 4:** Once the optimization process converges, evaluate the obtained solution and assess the FAR, FD accuracy, and other relevant performance metrics using the optimized decision variables, $\mathcal{L}^*$.

By following these steps and utilizing the appropriate mathematical expressions, SQP can effectively minimize false alarms and enhance fault detection accuracy, leading to improved system performance and operational safety. The parameter setting of SQP for MATLAB is shown in Table 2.

VI. CASE STUDIES

A. TENNESSEE EASTMAN PROCESS (TEP)

The Tennessee Eastman Process (TEP) represents a nonlinear, open-loop, and inherently unstable industrial process that has found wide application in various fields such as sensor fault detection, statistical process monitoring, and data-driven network study identification. Serving as a benchmark for industrial process analysis, TEP consists of five distinct processing units, each playing a crucial role in the overall operation:

1) **Two-Phase Reactor:** This unit serves as the core of the TEP, facilitating an exothermic reaction process.
2) **Separator:** The separator unit is responsible for separating the different components or phases generated during the reaction within the two-phase reactor.
3) **Stripper:** The stripper unit plays a critical role in the purification process, removing impurities or unwanted components from the mixture.
4) **Compressor:** The compressor unit is responsible for increasing the pressure of the process, enabling efficient transportation and further processing.
5) **Mixer:** The mixer unit combines different components or streams to achieve the desired composition or mixture.

These five processing units collectively form the TEP, representing a complex and interconnected system that poses challenges for analysis, control, and fault detection. Due to its inherent complexity, the TEP has emerged as a benchmark for evaluating the performance of various methodologies and techniques in the context of industrial process analysis and control. In this work, we demonstrate the efficacy of above mentioned data-driven approaches for computing FDR, ADR and MDR. The TEP involves a series of interconnected units

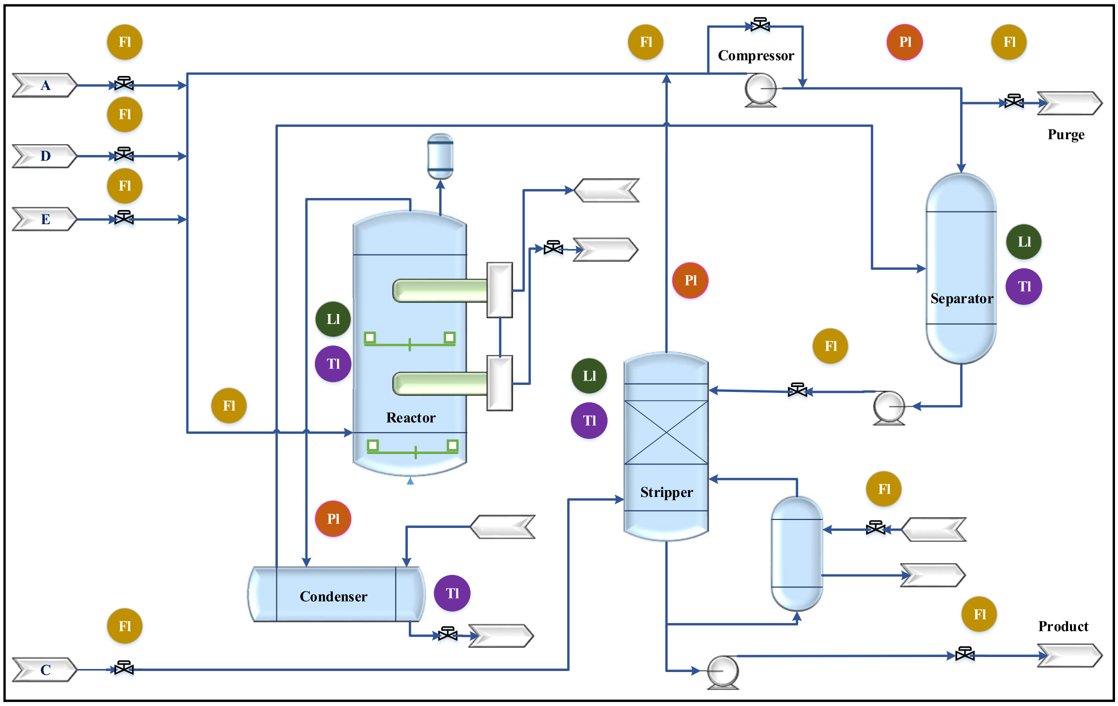


FIGURE 1. Schematic diagram of TEP.

TABLE 1. Comparison of methods with previous literature for TEP.

Sr. No	Parameters	Russel et al. [57]	Proposed work	Ji Lui [58]	Proposed work
		PCA		FDA	
		T2	T2	T2	T2
1	FAR (%)	0.62	1.4	3.1	0
2	MDR (%)	2.9	0	0.15	0.1
3	FDR (%)	95.35	97.14	97.14	93.15
4	TD (%)	1	1	0	3

that facilitate the desired chemical transformations. Initially, a reactor initiates the conversion of the gaseous feed components (A, C, D, and E) into liquid products (G and H). Subsequently, a condenser cools the gaseous product stream emerging from the reactor. Following this, a gas-liquid separator effectively segregates the gas and liquid components within the cooled product stream. To maintain the process continuity, a centrifugal compressor re-introduces the separated gas stream back into the reactor. Finally, a stripper unit enables the conversion of the separated gas and liquid components, fulfilling specific process requirements. The intricate interplay between these chemical elements and the series of units involved in the TEP exemplifies its complex nature. Understanding and effectively controlling these transformations and interactions within the process is essential for optimizing its performance and ensuring desired product outcomes.

The data consist of 52 variables and 960 samples while there are 12 manipulated variables, 41 measured variables and 21 faults of TEP. Normal data is the one with no faults. Figure 1 shows the diagram of the TEP.

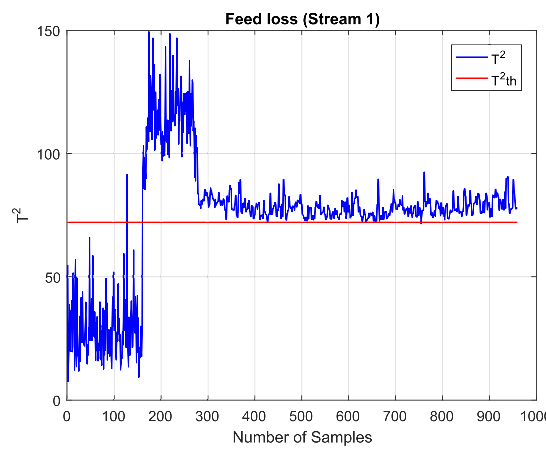


FIGURE 2. Fault detection of feed loss (Stream 1) fault using PCA for TEP.

r There are four different performance indicators for FDD for any industrial system which are False Alarm Rate (FAR), Missed Detection Rate (MDR), Fault Detection Rate (FDR)

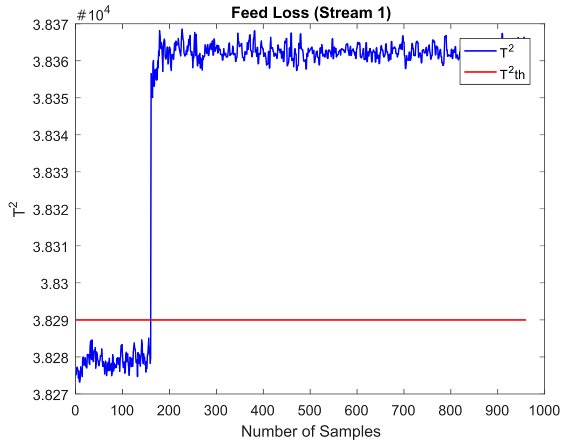


FIGURE 3. Fault detection of feedloss (Stream 1) fault using FDA for TEP.

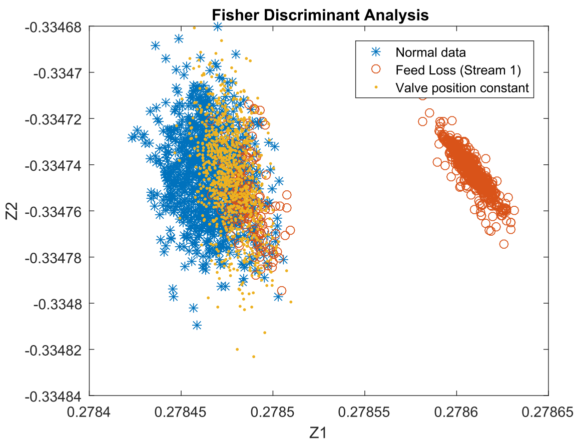


FIGURE 4. Fault isolation using FDA for TEP.

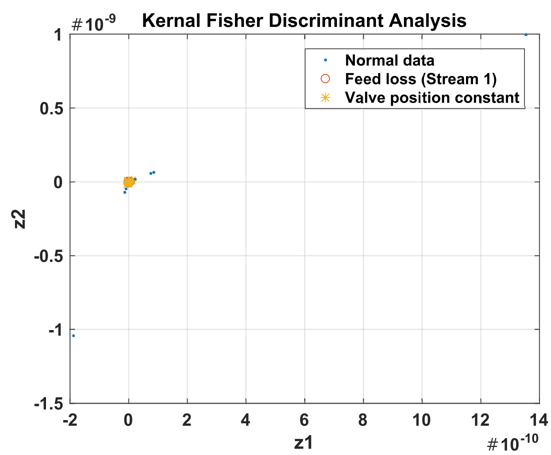


FIGURE 5. Fault isolation using KFDA for TEP.

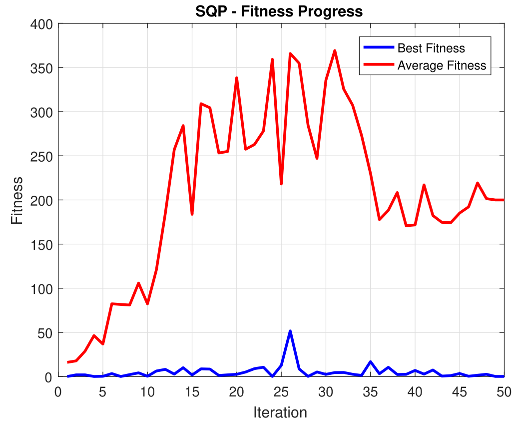


FIGURE 6. SQP convergence for TEP.

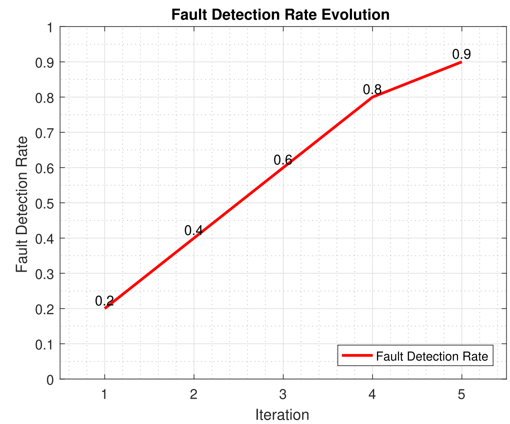


FIGURE 7. FDR by SQP for TEP.

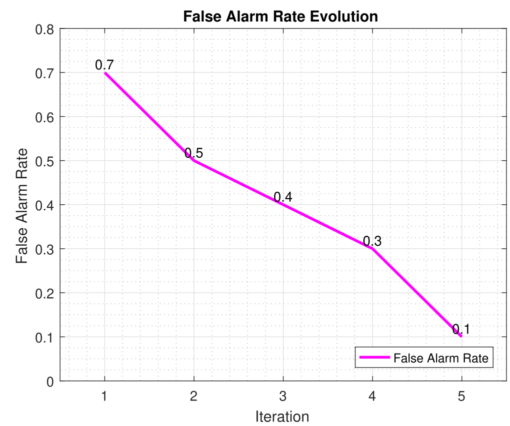


FIGURE 8. FAR by SQP for TEP.

and Time delay. FAR, MDR, FDR and TD are defined by following mathematical expressions [56],

$$FAR = \frac{\text{Samples above the limits}}{\text{Total number of normal samples}} \tag{35}$$

$$MDR = \frac{\text{faulty samples under the limits}}{\text{Total number of faulty samples}} \tag{36}$$

$$FDR = \frac{\text{Samples above the limits}}{\text{Total Number of faulty samples}} \tag{37}$$

$$TD = T_d - T_i \tag{38}$$

where T_d is the fault detected at time index while T_i is the fault introduced at time index. Table 1 shows the comparison

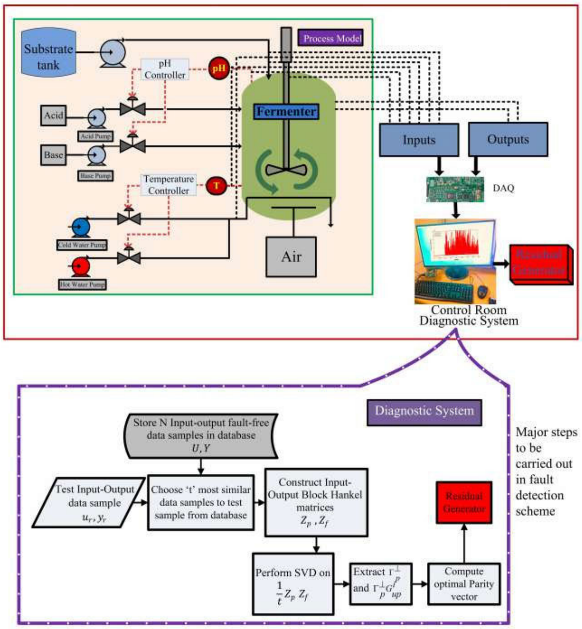


FIGURE 9. Schematic diagram of Penicillin Fermentation Process [61].

of False Alarm Rate (FAR), Missed Detection Rate (MDR), False Detection Rate (FDR) and Time Delay (TD) for PCA and FDA for TEP. It is observed that FAR is 0.62% using PCA while it is reduced to 0% using FDA. MDR is 0% in PCA while it is approximately 0.12% for FDA. FDR and TD are approximately same in PCA and FDA. TEP is implemented for PCA, FDA and KFDA for Fault Detection and Diagnosis (FDD). Figure 2 shows the fault detection for feedloss fault of TEP process using T^2 statistics for PCA. Figure 2 shows that the occurrence of a fault is at 160 samples while the maximum magnitude of T^2 is about 150. Fault detection of feedloss fault using FDA for TEP is shown in Figure 3. Figure 3 shows there is no false alarm while the fault detection rate is about 99.87%. Fault isolation between normal and different faults of TEP using FDA is shown in Figure 4. Similarly, fault isolation between normal and different faults of TEP using KFDA is shown in Figure 5. Comparison of figure 4 and 5 shows that data points of faults are very much isolated in KFDA as compared to FDA.

The convergence plot for the best and average fitness of SQP is presented in Figure 6, showcasing the rapid convergence of the proposed SQP optimization function towards the desired optimal point for the benchmark TEP. This convergence behavior demonstrates the efficacy of SQP in efficiently navigating the complex search space of the TEP, enabling the identification of optimal solutions. Figure 7 displays the FDR, providing empirical evidence of SQP's remarkable fault detection capabilities. The depicted FDR highlights the increased rate at which SQP accurately detects faults within the TEP, underscoring the effectiveness of the optimization-based approach in fault detection and diagnosis. Moreover, Figure 8 visually represents the successful minimization of the FAR achieved by SQP. The plot illustrates the ability of SQP to significantly reduce the occurrence of false alarms, which is crucial for enhancing the reliability and efficiency of fault detection systems. These figures collectively demonstrate the prowess of SQP in fault detection and alarm minimization for the TEP, emphasizing its utility as a powerful optimization tool in industrial processes.

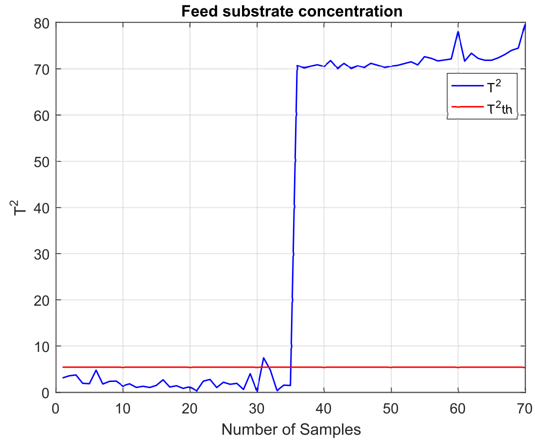


FIGURE 10. Fault detection of substrate concentration fault using PCA.

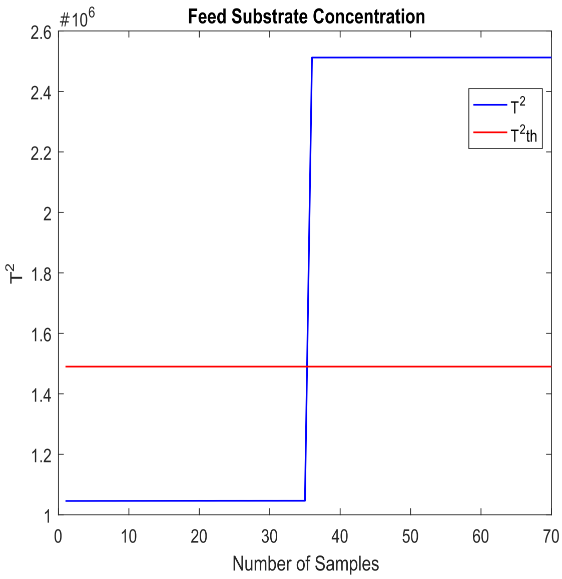


FIGURE 11. Fault detection of substrate concentration fault using FDA.

TABLE 2. Parametric adjustment of SQP.

Parameter	Value
Maximum Allowable Iterations	50
Maximum Fitness Assessment	250
TolFun	1e-12
TolCon	1e-10
TolX	1e-6
Others	Default

B. PENICILLIN FERMENTATION PROCESS (PFP)

Alexander Fleming discovers penicillin in 1928, making it one of the most often prescribed medicines [59]. Penicillin's fermentation procedure is an example of a nonlinear, dynamic

TABLE 3. Comparison of methods with previous work for PFP.

Sr. No	Parameters	Russel et al. [57]	Proposed work	Ji Lui [58]	Proposed work
		PCA		FDA	
		T2	T2	T2	T2
1	FAR (%)	1.4	0	0	0
2	MDR (%)	2.9	0	0.9	0
3	FDR (%)	95.35	99.87	97.14	99.15
4	TD (Samples)	1	1	0	1

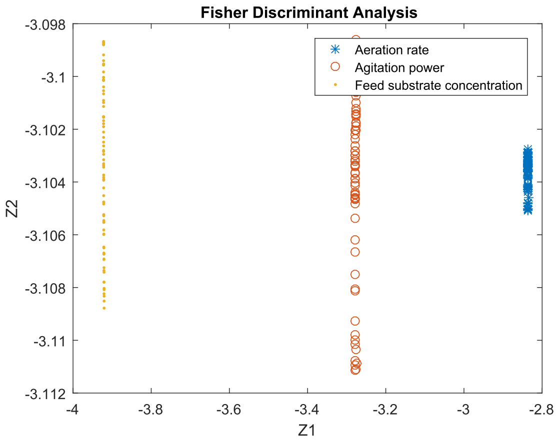


FIGURE 12. Fault isolation by FDA.

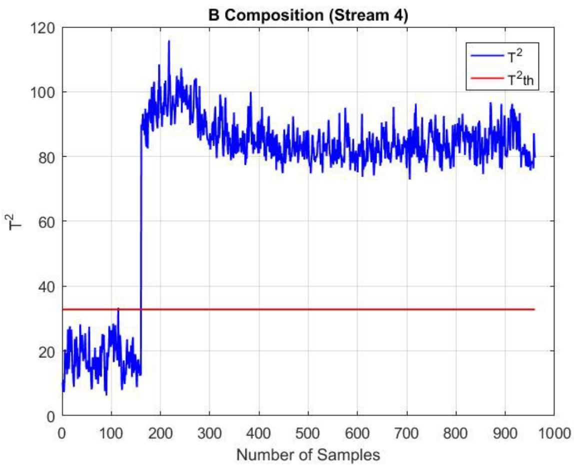


FIGURE 13. T^2 statistics is shown for B Composition (Stream 4) fault using PCA for TEP.

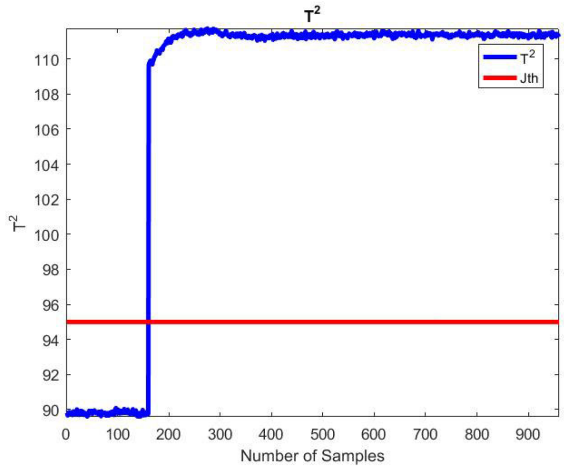


FIGURE 14. T^2 statistics is shown for B Composition (Stream 4) fault using FDA for TEP.

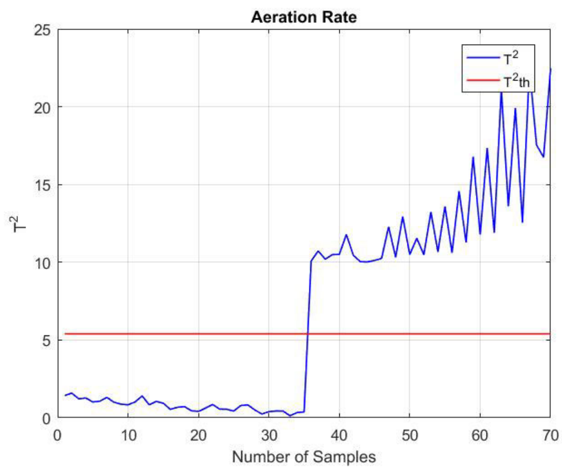


FIGURE 15. Aeration rate fault is simulated using PCA for PFP.

batch process. Two stages comprise industrial penicillin production: the pre-culture stage and the fed-batch stage. First of all, penicillin consumes a significant percentage of the initial substrate supply, causing the substrate to decrease. Continuous maintenance of the substrate is a related open-loop process [60]. In the second step, penicillin is enhanced and released into the medium. When performed at an acidic pH, the overall efficiency of absorption increases by 2 to 5%. Butyl, amyl, or isobutyl acetate is utilized to extract the substrate from a cold acidic soap. Fermentation processes for penicillin are now highly automated. During the active production phase, the pH is maintained between 5.5 and 6, but it can reach 7 due to the consumption of lactic acid or the release of NH_3. $MgCO_3$, $CaCO_3$, or phosphate buffer will be added if the pH is higher than 8. Agitator power is 30W, and the rate of aeration ranges from 30 to 60 L/h, which is initially high and then produces less oxygen over time [60]. Thus, an improvement in PFP's safety and dependability is highly desired. Errors must be minimized and prevented if penicillin production is to continue at a high level of quality. This research endeavor is not primarily concerned with the precise explanations of biosynthesis and chemical processes. Chemical reactions occurred in a process are followed from the literature [60]. This model disregards environmental

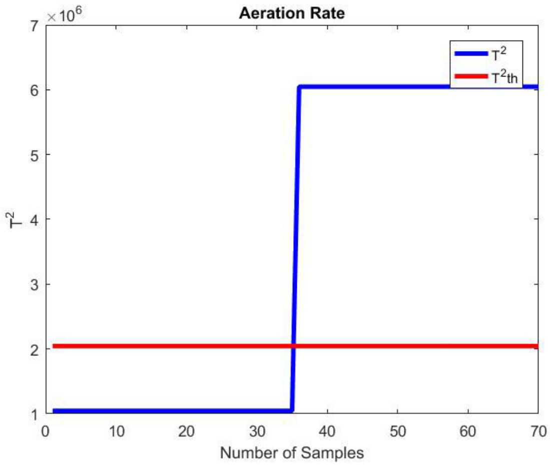


FIGURE 16. Aeration rate fault is simulated using FDA for PFP.

variables such as temperature and pH. The simplified diagram of PFP is shown in Figure 9.

The training data consist of 6 measured variables and 70 samples of while the testing data consist of 6 measured variables and 50 samples. There are 16 faults of PFP. Figure 10 shows the fault detection of feed substrate concentration fault using PCA. Table 3 shows that FAR is 2.85% using PCA while it is reduced to 0% for FDA. Similarly, MDR is 0% using PCA while it 2.87% using FDA. A plot of detection for FDA is shown in Figure 11. Figure 12 shows the faults isolation by FDA while fault isolation performed by KFDA is shown in Figure 5. It is observed that isolation is much better achieved in KFDA as compared to FDA.

To support the claims, further nonlinear valve data is simulated using the offered methodologies. Figure 13 displays the T^2 statistics for the defect of B Composition (Stream 4) using PCA for TEP. The detection of a problem is seen to occur at the 160th sample. The data demonstrates that both the FAR and the MDR are recorded as 0%, indicating a high level of reliability and accuracy in the obtained results. Figure 14 presents the T^2 statistics for the fault in B Composition (Stream 4) using the FDA for TEP. The findings confirm that the FAR and the MDR both register at 0% when employing the FDA methodology. Figure 15 depicts the simulation of the PFP PCA. The simulation results indicate that the FAR is 1.3%, but the MDR is 0.5%. The simulation of PFP system is conducted using the FDA, as seen in Figure 16. The simulation results indicate that the FAR is 0.5%, and the MDR is 0.10%.

C. RESEARCH SUMMARY AND MANAGERIAL INSIGHTS: ENHANCING INDUSTRIAL FAULT DETECTION AND ISOLATION

In summary, our research has conducted a comprehensive examination of three data-driven design methodologies, namely PCA, FDA, and KFDA. Additionally, we have explored the use of SQP optimization for the purpose of detecting and isolating faults in industrial settings. The utilization of complex mathematical principles in these systems entails intricate technical details, however their consequences for industrial management are of great significance. The results of our study highlight the possibility of improving problem diagnostics in intricate industrial systems, leading to reduced operational interruptions and increased system dependability. The following insights offer useful implications for managers and decision-makers who may have limited familiarity with the underlying mathematical concepts:

1) **Tailored Approach Selection:** Each approach, namely PCA, FDA, and KFDA, have distinct advantages. PCA and FDA have been found to be effective in detecting faults, making them accessible to non-technical readers. On the other hand, KFDA has been recognized for its exceptional ability to isolate faults, distinguishing it from the other methods. This observation empowers managers to effectively synchronize their choice of methods with their operational requirements.
2) **Trade-off Between Alarms and Detection:** Managers possess the capacity to recognize and comprehend the inherent compromise between the occurrence of FAR and the potential for MDR. While the Fisher Discriminant Analysis (FDA) is effective in reducing false alarms when compared to PCA, it is possible that the FDA may exhibit a little greater percentage of missed detections. This trade-off has an impact on the manner in which operational teams choose the order of priority for addressing alerts.
3) **Resource Optimization with KFDA:** The significance of KFDA's fault isolation proficiency becomes evident to readers who possess less knowledge in mathematical models. The exact fault localization capability of the KFDA enables managers to effectively allocate resources for maintenance activities, hence minimizing downtime.
4) **Precision through SQP Optimization:** The use of SQP optimization techniques has substantial administrative ramifications. Non-specialized readers may comprehend that this improvement contributes to the reduction of false alarms, enabling operators to rapidly prioritize significant concerns and improve overall efficiency.

Fundamentally, although the study incorporates intricate mathematical models, the key insights for managers and decision-makers are practical. Managers may make educated decisions on fault detection and isolation procedures by comprehending the practical implications discussed earlier. This will result in optimized operations, reduced downtime, and enhanced system dependability.

D. ASSOCIATED ASSUMPTIONS AND LIMITATIONS OF COMPARATIVE STUDY

PCA, FDA, and KFDA are dimensionality reduction techniques commonly used in various fields, including industrial fault detection and diagnosis. When applying these, there are

TABLE 4. Statistical test for comparative methods.

Sr. No	Methods	Case Studies							
		Tenessee Eastman Process				Penicillin Fermentation Process			
		Mean	Min	Max	STD	Mean	Min	Max	STD
1	PCA	4.31	4.10	4.89	4.71	2.94	2.45	2.99	2.93
2	FDA	4.24	4.13	5.63	4.41	2.8378e-05	2.8336e-05	2.8407e-5	2.125
3	KFDA	2.38405e-06	2.2375e-06	2.4875e-06	2.21	2.4875e-06	2.4870e-06	2.4879e-06	2.11
4	SQP	3.434e+07	2.635e+07	4.319e+07	2.940e+06	2.402e+03	1.711e+03	3.242e+03	2.854e+04

TABLE 5. Complexity comparison: n is the number of samples, d is the dimension, p and q are the number of original testing samples, t is the rank and c is the number of classes.

Sr. No.	Methods	Computational complexity	Run time of TEP	Run time of PFP
1	PCA	O$(d^2hL + d^n)$ [68]	1.74	0.60
2	FDA	O$(c+p+q)n^2 + (c+p+q)^2 n)$ [69]	1.35	2.84
3	KFDA	O$(n^3 + t(d^3 + nd^2 + n^2 d))$ [70]	1.61	1.55
4	SQP	O$(n^2 + nd^2 + n^2 d)$ [71]	3.01	2.98

TABLE 6. Comparison of various FDD approaches in industrial control.

Reference	Features	Observer	Digraphs	Abstraction hierarchy	Smart Systems	PCA	FDA	KFDA	SQP
[72]	Quick FDD	✓	×	✓	✓	✓	✓	✓	×
[73]	Fault separability	✓	×	×	×	✓	✓	×	×
[74]	Robustness	×	✓	✓	×	×	×	✓	✓
[75]	Distinctiveness discernment	✓	×	×	✓	✓	×	×	✓
[76]	Misclassification rate	×	×	×	×	×	×	✓	✓
[77]	Flexibility	×	✓	✓	×	×	×	×	✓
[78]	Multiplicity detection	×	×	×	✓	×	×	✓	✓

several necessary assumptions and considerations to keep in mind:

1) PCA: Assumptions are essential when using PCA for FDD in TEP and PEP. PCA presupposes linearity and statistical independence, which may work in present case studies, but avoid nonlinearities. Standardize variables to eliminate scaling difficulties and assure normal distribution for reliable findings. Outliers should be addressed to avoid skewing results. Avoid overfitting by having more observations than variables and interpreting industrial system primary components meaningfully. Balance variability preservation and dimensionality reduction while selecting preserved components. Use domain knowledge to connect mathematical and physical interpretations. If data is temporal, try Dynamic PCA. Validate and update the PCA model to reflect changing operational circumstances. PCA works with other methodologies and expert opinions to provide a complete industrial strategy [62].
2) FDA: Using FDA procedures in industry relies on fundamental assumptions to ensure process effectiveness and dependability. Deviations from a steady, well-defined operating behavior indicate fault. A detailed knowledge of the system's behavior of case studies and dynamics are needed to define normal functioning and identify deviations. It is assumed that fault isolation allows for diverse reactions to deviations. An assumption of a largely constant operational environment is established since rapid changes may affect sensor accuracy. Historical data is assumed to enable model creation and performance benchmarking. These principles drive FDA application, but their relevance and adaptability should be assessed based on unique procedures and industrial circumstances [63].
3) KFDA: The kernel, a key part of KFDA, affects data translation into higher dimensions. Additionally, KFDA works best when classes are clearly separated. This assumption supports the view that the modified feature space well-separates fault situations. A small dataset may lead to incorrect sample estimates and poor model performance. Resampling or class weighted KFDA may be needed due to imbalanced class distributions [64].

It is important to note that these strategies include inherent limits and are based on certain assumptions. In the field of industrial defect detection and diagnosis, the integration of these techniques with other approaches, such as domain-specific information and additional multivariate statistical methods, has the potential to enhance the reliability and precision of the constructed models. We carried out statistical tests for the T^2 analysis of feed loss (stream 1) for TEP and substrate concentration for the PFP process. The statistical results in terms of mean, maximum, minimum, and standard deviation are summarized in Table 4. The comparative analysis shows that residual signal for KFDA is better than PCA and FDA as the standard deviation is less so results are more accurate for fault isolation. Similarly, the mean is less in the case of KFDA which shows that the error signal is much more accurate.

E. CONTROLLING PARAMETERS

The main performance indexes for controlling the proposed method are FAR, MDR, FDR and TD which should be up to the appropriate level [65]. For the selection of optimal controlling parameters in PCA for threshold computation confidence interval of F-distribution is set to 99%. Moreover, c_β is set between -5 to 10. In addition, a number of principal components (PC's) are 16 which controls the value of T^2 and T^2 threshold statistics. It should be noted that the criteria for selecting the number of PCs are on the basis of Cumulative Percentage Variance (CPV) [66]. In FDA, major principal components are fewer than PCA based on the solution of generalized eigenvector solution. KFDA includes variable kernel functions such as Polyplus, linear, Gaussian, RBK kernel, Laplacian kernel etc. which take dot product with eigenvectors in the higher dimension and cause isolation of faults by use of various transformation matrices.

The comparative methods are robust as faults are detected by maintaining a FAR less than 2% [67]. Table 1 shows that TEP and PFP case studies with T^2 statistics simulations show that FAR is lower than 2 % which shows that the system is able to withstand uncertainties in the presence of faults. Moreover, Tables 1 and 3 show the comparison of our result with other research work present in the literature. The computational complexity of approaches for TEP process and PFP process is summarized in Table 5. Moreover, run times for both case studies are shown in Table 5. It is noted by a comparative analysis that PCA is recommendable considering its low computational complexity while KFDA has the highest computational complexity. To sum up, it is observed that the cost-effectiveness of PCA is higher as compared to other techniques.

In this article, four techniques are compared for fault diagnosis including fault detection, fault isolation, and fault identification. To make the comparisons fair, we let the numbers of nonzero loading vectors (PC's) same among the four methods and compare the differences in transformation vectors for different methods. It is observed that the results of simulation for detection and isolation are different due to the difference in the transformation matrix.

VII. CONCLUSION

In conclusion, this study has provided a comprehensive analysis of three data-driven design schemes, namely PCA, FDA, KFDA, and SQP for fault detection and isolation in industrial applications. Our findings highlight the strengths and limitations of each method, shedding light on their respective advantages and areas for improvement. Both PCA and FDA demonstrate their effectiveness in fault detection, while KFDA excels in fault isolation. Specifically, FDA exhibits a notable reduction in the FAR compared to PCA. However, it is important to note that FDA also shows a slightly higher MDR in comparison. On the other hand, KFDA receives high acclaim for its fault isolation approach, showcasing its potential for accurately pinpointing faults in industrial processes. Additionally, we observe that both PCA and FDA achieve FDR exceeding 97%, underscoring their robustness and reliability in detecting faults. These results highlight the efficacy of data-driven design schemes in enhancing fault diagnosis in complex industrial systems. Building upon these findings, several future directions can be explored. Firstly, further research can focus on refining FDA to strike a better balance between FAR and MDR, potentially through the incorporation of advanced classification techniques or algorithmic enhancements. Secondly, investigations can be pursued to optimize KFDA for fault detection, potentially by leveraging the strengths of other machine learning algorithms or exploring novel feature extraction methods. By leveraging the power of mathematical optimization, this approach enables efficient and accurate fault detection, leading to reduced downtime and increased system reliability. The use of SQP optimization also aids in minimizing false alarms, ensuring that operators can focus their attention on critical issues, enhancing overall operational efficiency. Overall, SQP optimization proves to be a valuable tool in optimizing fault detection and alarm systems as compared to FDA and KFDA, contributing to improved performance and maintenance in TEP.

Moreover, future studies may explore the integration of multiple data-driven design schemes to harness the collective power of different methods, leading to improved overall fault detection and isolation performance. Additionally, the application of these schemes to different industrial processes and the utilization of real-time data can be further investigated to enhance the accuracy and timeliness of fault diagnosis.

REFERENCES

[1] A. Basit, M. Tufail, M. Rehan, and I. Ahmed, "A new event-triggered distributed state estimation approach for one-sided Lipschitz nonlinear discrete-time systems and its application to wireless sensor networks," *ISA Trans.*, vol. 137, pp. 74–86, Jun. 2023.

[2] U. Akram, M. Khalid, and S. Shafiq, "An improved optimal sizing methodology for future autonomous residential smart power systems," *IEEE Access*, vol. 6, pp. 5986–6000, 2018.

[3] M. Khalid, U. Akram, and S. Shafiq, "Optimal planning of multiple distributed generating units and storage in active distribution networks," *IEEE Access*, vol. 6, pp. 55234–55244, 2018.

[4] I. Ahmed, M. Rehan, A. Basit, M. Tufail, and K.-S. Hong, "A dynamic optimal scheduling strategy for multi-charging scenarios of plug-in-electric vehicles over a smart grid," *IEEE Access*, vol. 11, pp. 28992–29008, 2023.

[5] S. Alshahrani, M. Khalid, and M. Almuhaini, "Electric vehicles beyond energy storage and modern power networks: Challenges and applications," *IEEE Access*, vol. 7, pp. 99031–99064, 2019.

[6] F. E. Mustafa, I. Ahmed, A. Basit, U.-E.-H. Alvi, S. H. Malik, A. Mahmood, and P. R. Ali, "A review on effective alarm management systems for industrial process control: Barriers and opportunities," *Int. J. Crit. Infrastruct. Protection*, vol. 41, Jul. 2023, Art. no. 100599.

[7] Z. Deng, X. Chen, Y. Xie, Z. Zou, and H. Zhang, "Multiple structured latent double dictionary pair learning for cross-domain industrial process monitoring," *Inf. Sci.*, vol. 648, Nov. 2023, Art. no. 119514.
[8] J. Yang, F. Cheng, Z. Liu, M. M. Duodu, and M. Zhang, "A novel semi-supervised fault detection and isolation method for battery system of electric vehicles," *Appl. Energy*, vol. 349, Nov. 2023, Art. no. 121650.
[9] S. Zhang, L. Su, J. Gu, K. Li, L. Zhou, and M. Pecht, "Rotating machinery fault detection and diagnosis based on deep domain adaptation: A survey," *Chin. J. Aeronaut.*, vol. 36, no. 1, pp. 45–74, Jan. 2023.
[10] S. Marguet, "The three-mile-island accident," in *A Brief History of Nuclear Reactor Accidents: From Leipzig to Fukushima*. Cham, Switzerland: Springer, 2023, pp. 211–265.
[11] F. Zhang, N. Saeed, and P. Sadeghian, "Deep learning in fault detection and diagnosis of building HVAC systems: A systematic review with meta analysis," *Energy AI*, vol. 12, Apr. 2023, Art. no. 100235.
[12] J. Chu, N. Sun, W. Hu, X. Chen, N. Yi, and Y. Shen, "The application of Bayesian methods in cancer prognosis and prediction," *Cancer Genomics-Proteomics*, vol. 19, no. 1, pp. 1–11, 2022.
[13] L. Elkhidir, A. Hassan, and M. Khalid, "SVC-based controller design via ant colony optimization algorithm," in *Proc. 8th Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Nov. 2019, pp. 301–308.
[14] U.-E.-H. Alvi, I. Ahmed, S. R. Hasan, B. Ashfaq, M. Raza, and S. Mukhtar, "Adaptive swarm intelligence-based optimization approach for smart grids power dispatch," in *Proc. Int. Conf. Emerg. Technol. Electron., Comput. Commun. (ICETECC)*, Dec. 2022, pp. 1–6.
[15] S. Alshahrani, M. Al-Muhaini, and M. Khalid, "Minimizing active/reactive power losses in electricity networks based on optimal location of battery energy storage system," in *Proc. 8th Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Nov. 2019, pp. 287–294.
[16] I. Ahmed, M. Rehan, A. Basit, M. Tufail, and K.-S. Hong, "Neuro-fuzzy and networks-based data driven model for multi-charging scenarios of plug-in-electric vehicles," *IEEE Access*, vol. 11, pp. 87150–87165, 2023.
[17] I. Ahmed, M. Rehan, K.-S. Hong, and A. Basit, "Event-triggered leaderless robust consensus control of nonlinear multi-agents under disturbances," in *Proc. 13th Asian Control Conf. (ASCC)*, May 2022, pp. 1736–1741.
[18] U.-E.-H. Alvi, I. Ahmed, A. Alvi, B. Ashfaq, S. Mukhtar, and P. R. Ali, "Technological, financial and ecological analysis of photovoltaic power system using RETScreen®: A case in Khuzdar, Pakistan," in *Proc. Int. Conf. Emerg. Technol. Electron., Comput. Commun. (ICETECC)*, Dec. 2022, pp. 1–6.
[19] D. C. Caixeta, C. Lima, Y. Xu, M. Guevara-Vega, F. S. Espindola, R. Goodacre, D. M. Zezell, and R. Sabino-Silva, "Monitoring glucose levels in urine using FTIR spectroscopy combined with univariate and multivariate statistical methods," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 290, Apr. 2023, Art. no. 122259.
[20] M. Khalid, R. P. Aguilera, A. V. Savkin, and V. G. Agelidis, "On maximizing profit of wind-battery supported power station based on wind power and energy price forecasting," *Appl. Energy*, vol. 211, pp. 764–773, Feb. 2018.
[21] M. Khalid and A. V. Savkin, "Minimization and control of battery energy storage for wind power smoothing: Aggregated, distributed and semi-distributed storage," *Renew. Energy*, vol. 64, pp. 105–112, Apr. 2014.
[22] I. Jolliffe, "A 50-year personal journey through time with principal component analysis," *J. Multivariate Anal.*, vol. 188, Mar. 2022, Art. no. 104820.
[23] J. Barbier, F. Camilli, M. Mondelli, and M. Sáenz, "Fundamental limits in structured principal component analysis and how to reach them," *Proc. Nat. Acad. Sci. USA*, vol. 120, no. 30, Jul. 2023, Art. no. e2302028120.
[24] F. Zhu, J. Gao, J. Yang, and N. Ye, "Neighborhood linear discriminant analysis," *Pattern Recognit.*, vol. 123, Mar. 2022, Art. no. 108422.
[25] X. Bai, M. Zhang, Z. Jin, Y. You, and C. Liang, "Fault detection and diagnosis for chiller based on feature-recognition model and kernel discriminant analysis," *Sustain. Cities Soc.*, vol. 79, Apr. 2022, Art. no. 103708.
[26] X. Qin, T. R. Lock, and R. L. Kallenbach, "DA: Population structure inference using discriminant analysis," *Methods Ecol. Evol.*, vol. 13, no. 2, pp. 485–499, Feb. 2022.
[27] A. Atif and M. Khalid, "Saviztky–Golay filtering for solar power smoothing and ramp rate reduction based on controlled battery energy storage," *IEEE Access*, vol. 8, pp. 33806–33817, 2020.
[28] M. Li, P. Zhang, G. Yang, G. Xu, M. Guo, and W. Liao, "A Fisher linear discriminant analysis classifier fused with naïve Bayes for simultaneous detection in an asynchronous brain–computer interface," *J. Neurosci. Methods*, vol. 371, Apr. 2022, Art. no. 109496.
[29] A. Ozier-Lafontaine, C. Fourneaux, G. Durif, C. Vallot, O. Gandrillon, S. Giraud, B. Michel, and F. Picard, "Kernel-based testing for single-cell differential analysis," 2023, *arXiv:2307.08509*.
[30] Q. Liu, X. Li, H. Liu, and Z. Guo, "Multi-objective metaheuristics for discrete optimization problems: A review of the state-of-the-art," *Appl. Soft Comput.*, vol. 93, Aug. 2020, Art. no. 106382.
[31] K. Ransikarbum and S. J. Mason, "A bi-objective optimisation of post-disaster relief distribution and short-term network restoration using hybrid NSGA-II algorithm," *Int. J. Prod. Res.*, vol. 60, no. 19, pp. 5769–5793, Oct. 2022.
[32] I. Ahmed, M. Rehan, A. Basit, and K.-S. Hong, "Greenhouse gases emission reduction for electric power generation sector by efficient dispatching of thermal plants integrated with renewable systems," *Sci. Rep.*, vol. 12, no. 1, p. 12380, Jul. 2022.
[33] I. Ahmed, U.-E.-H. Alvi, A. Basit, M. Rehan, and K.-S. Hong, "Multi-objective whale optimization approach for cost and emissions scheduling of thermal plants in energy hubs," *Energy Rep.*, vol. 8, pp. 9158–9174, Nov. 2022.
[34] U. Akram, M. Khalid, and S. Shafiq, "Optimal sizing of a wind/solar/battery hybrid grid-connected microgrid system," *IET Renew. Power Gener.*, vol. 12, no. 1, pp. 72–80, Jan. 2018.
[35] I. Ahmed, M. Rehan, A. Basit, S. H. Malik, U.-E.-H. Alvi, and K.-S. Hong, "Multi-area economic emission dispatch for large-scale multi-fueled power plants contemplating inter-connected grid tie-lines power flow limitations," *Energy*, vol. 261, Dec. 2022, Art. no. 125178.
[36] S. Na, M. Anitescu, and M. Kolar, "An adaptive stochastic sequential quadratic programming with differentiable exact augmented Lagrangians," *Math. Program.*, vol. 199, nos. 1–2, pp. 721–791, May 2023.
[37] H. Jiang, Z. Li, Y. Sun, S. Jiang, and J. Tian, "Optimization of the countercurrent continuous reforming process based on equation-oriented modeling and the SQP algorithm," *ACS Omega*, vol. 7, no. 2, pp. 1757–1771, Jan. 2022.
[38] F. Yao, Z. Chen, X. Liu, and M. Zhang, "Weak thruster fault detection for AUV based on Bayesian network and hidden Markov model," *Proc. Inst. Mech. Eng., M, J. Eng. Maritime Environ.*, vol. 237, no. 2, pp. 478–486, May 2023.
[39] Y. Han, N. Ding, Z. Geng, Z. Wang, and C. Chu, "An optimized long short-term memory network based fault diagnosis model for chemical processes," *J. Process Control*, vol. 92, pp. 161–168, Aug. 2020.
[40] S. R. K. Priya, R. K. Balambiga, P. Mishra, and S. S. Das, "Sugarcane yield forecast using weather based discriminant analysis," *Smart Agricult. Technol.*, vol. 3, Feb. 2023, Art. no. 100076.
[41] A. Kaul and S. Raina, "Support vector machine versus convolutional neural network for hyperspectral image classification: A systematic review," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 15, p. e6945, Jul. 2022.
[42] E. O. Omuya, G. O. Okeyo, and M. W. Kimwele, "Feature selection for classification using principal component analysis and information gain," *Expert Syst. Appl.*, vol. 174, Jul. 2021, Art. no. 114765.
[43] J. Liu, H. Kang, W. Tao, H. Li, D. He, L. Ma, H. Tang, S. Wu, K. Yang, and X. Li, "A spatial distribution—Principal component analysis (SD-PCA) model to assess pollution of heavy metals in soil," *Sci. Total Environ.*, vol. 859, Feb. 2023, Art. no. 160112.
[44] R. J. Rainer, M. Mayr, J. Himmelbauer, and R. Nikzad-Langerodi, "Opening the black-box of neighbor embeddings with Hotelling's T^2 statistic and Q-residuals," *Chemometric Intell. Lab. Syst.*, vol. 238, Jul. 2023, Art. no. 104840.
[45] B. M. S. Hasan and A. M. Abdulazeez, "A review of principal component analysis algorithm for dimensionality reduction," *J. Soft Comput. Data Mining*, vol. 2, no. 1, pp. 20–30, 2021.
[46] J. R. Beattie and F. W. L. Esmonde-White, "Exploration of principal component analysis: Deriving principal component analysis visually using spectra," *Appl. Spectrosc.*, vol. 75, no. 4, pp. 361–375, Apr. 2021.
[47] J. Pino-Ortega, D. Rojas-Valverde, C. D. Gómez-Carmona, and M. Rico-González, "Training design, performance analysis, and talent identification—A systematic review about the most relevant variables through the principal component analysis in soccer, basketball, and rugby," *Int. J. Environ. Res. Public Health*, vol. 18, no. 5, p. 2642, Mar. 2021.
[48] J. Qian, Y. Cao, Y. Bi, H. Wu, Y. Liu, Q. Chen, and C. Zuo, "Structured illumination microscopy based on principal component analysis," *eLight*, vol. 3, no. 1, p. 4, Feb. 2023.

[49] H. Wang, R. Liu, S. X. Ding, Q. Hu, Z. Li, and H. Zhou, "Causal-trivial attention graph neural network for fault diagnosis of complex industrial processes," *IEEE Trans. Ind. Informat.*, early access, Jun. 8, 2023, doi: 10.1109/TII.2023.3282979.

[50] W. Jia, Z. Zong, and T. Lan, "Elastic impedance inversion incorporating fusion initial model and kernel Fisher discriminant analysis approach," *J. Petroleum Sci. Eng.*, vol. 220, Jan. 2023, Art. no. 111235.

[51] S. Dong, L. Zeng, X. Du, J. He, and F. Sun, "Lithofacies identification in carbonate reservoirs by multiple kernel Fisher discriminant analysis using conventional well logs: A case study in a oilfield, Zagros basin, Iraq," *J. Petroleum Sci. Eng.*, vol. 210, Mar. 2022, Art. no. 110081.

[52] X. Tao, C. Ren, Q. Li, W. Guo, R. Liu, Q. He, and J. Zou, "Bearing defect diagnosis based on semi-supervised kernel local Fisher discriminant analysis using pseudo labels," *ISA Trans.*, vol. 110, pp. 394–412, Apr. 2021.

[53] F. Jamil, M. Abid, M. Adil, I. Haq, A. Q. Khan, and S. F. Khan, "Kernel approaches for fault detection and classification in PARR-2," *J. Process Control*, vol. 64, pp. 1–6, Apr. 2018.

[54] S. Mia, M. H. Talukder, and M. M. Rahman, "RobustDespeckling: Robust speckle noise reduction method using multi-scale and kernel Fisher discriminant analysis," *Biomed. Eng. Adv.*, vol. 5, Jun. 2023, Art. no. 100085.

[55] J. Liu, W. Xu, F. Zhang, and H. Lian, "Properties of standard and sketched kernel Fisher discriminant," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 10596–10602, Aug. 2023.

[56] M. Adil, M. Abid, A. Q. Khan, and G. Mustafa, "Comparison of PCA and FDA for monitoring of coupled liquid tank system," in *Proc. 13th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2016, pp. 225–230.

[57] E. L. Russell, L. H. Chiang, and R. D. Braatz, "Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis," *Chemometric Intell. Lab. Syst.*, vol. 51, no. 1, pp. 81–93, May 2000.

[58] B. Jiang, X. Zhu, D. Huang, J. A. Paulson, and R. D. Braatz, "A combined canonical variate analysis and Fisher discriminant analysis (CVA–FDA) approach for fault diagnosis," *Comput. Chem. Eng.*, vol. 77, pp. 1–9, Jun. 2015.

[59] L. Hua, C. Zhang, W. Sun, Y. Li, J. Xiong, and M. S. Nazir, "An evolutionary deep learning soft sensor model based on random forest feature selection technique for penicillin fermentation process," *ISA Trans.*, vol. 136, pp. 139–151, May 2023.

[60] L. Li, N. Li, X. Wang, J. Zhao, H. Zhang, and T. Jiao, "Multi-output soft sensor modeling approach for penicillin fermentation process based on features of big data," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 119208.

[61] M. A. Abbasi, A. Q. Khan, G. Mustafa, M. Abid, A. S. Khan, and N. Ullah, "Data-driven fault diagnostics for industrial processes: An application to penicillin fermentation process," *IEEE Access*, vol. 9, pp. 65977–65987, 2021.

[62] G. L. P. Palla and A. K. Pani, "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*, vol. 209, Mar. 2023, Art. no. 112504.

[63] E. Landsaat and J. Lipka, "Logistic regression and linear discriminant analysis: A comparative overview and an empirical time-complexity analysis," in *Proc. 20th SC@ RUG*, 2023, p. 59.

[64] A. Nasri, A. Boujnah, A. Boubaker, and A. Kalboussi, "A smart gas sensor using machine learning algorithms: Sensor types based on IED configurations, fabrication techniques, algorithmic approaches, challenges, progress and limitations: A review," *IEEE Sensors J.*, vol. 23, no. 11, pp. 11336–11355, Jun. 2023.

[65] W. Jung, S.-H. Kim, S.-H. Yun, J. Bae, and Y.-H. Park, "Vibration, acoustic, temperature, and motor current dataset of rotating machine under varying operating conditions for fault diagnosis," *Data Brief*, vol. 48, Jun. 2023, Art. no. 109049.

[66] M. Ammiche, A. Kouadri, and A. Bakdi, "A combined monitoring scheme with fuzzy logic filter for plant-wide Tennessee eastman process fault detection," *Chem. Eng. Sci.*, vol. 187, pp. 269–279, Sep. 2018.

[67] D. Wanner, A. Trigell, L. Drugge, and J. Jerrelind, "Survey on fault-tolerant vehicle design," *World Electric Vehicle J.*, vol. 5, no. 2, pp. 598–609, Jun. 2012.

[68] A. Sharma and K. K. Paliwal, "Fast principal component analysis using fixed-point algorithm," *Pattern Recognit. Lett.*, vol. 28, no. 10, pp. 1151–1155, Jul. 2007.

[69] H. Zhao and P. C. Yuen, "Incremental linear discriminant analysis for face recognition," *IEEE Trans. Syst., Man, Cybern., B, Cybernetics*, vol. 38, no. 1, pp. 210–221, Feb. 2008.

[70] D. Cai, X. He, and J. Han, "Speed up kernel discriminant analysis," *VLDB J.*, vol. 20, no. 1, pp. 21–33, Feb. 2011.

[71] I. Ahmed, U.-E.-H. Alvi, A. Basit, T. Khursheed, A. Alvi, K.-S. Hong, and M. Rehan, "A novel hybrid soft computing optimization framework for dynamic economic dispatch problem of complex non-convex contiguous constrained machines," *PLoS ONE*, vol. 17, no. 1, Jan. 2022, Art. no. e0261709.

[72] X. Bi, R. Qin, D. Wu, S. Zheng, and J. Zhao, "One step forward for smart chemical process fault detection and diagnosis," *Comput. Chem. Eng.*, vol. 164, Aug. 2022, Art. no. 107884.

[73] N. Md Nor, C. R. Che Hassan, and M. A. Hussain, "A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems," *Rev. Chem. Eng.*, vol. 36, no. 4, pp. 513–553, May 2020.

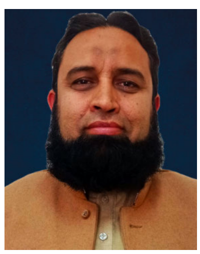
[74] N. Md Nor, M. A. Hussain, and C. R. Che Hassan, "Multi-scale kernel Fisher discriminant analysis with adaptive neuro-fuzzy inference system (ANFIS) in fault detection and diagnosis framework for chemical process systems," *Neural Comput. Appl.*, vol. 32, no. 13, pp. 9283–9297, Jul. 2020.

[75] M. Sani, M. Piffard, and V. Heiries, "Fault detection for PEM fuel cells via analytical redundancy: A critical review and prospects," *Energies*, vol. 16, no. 14, p. 5446, Jul. 2023.

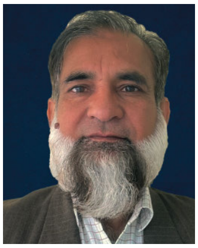
[76] A. Bradford, M. E. Kunik, P. Schulz, S. P. Williams, and H. Singh, "Missed and delayed diagnosis of dementia in primary care: Prevalence and contributing factors," *Alzheimer Disease Associated Disorders*, vol. 23, no. 4, pp. 306–314, 2009.

[77] J. Chen, L. Zhang, Y. Li, Y. Shi, X. Gao, and Y. Hu, "A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems," *Renew. Sustain. Energy Rev.*, vol. 161, Jun. 2022, Art. no. 112395.

FAIZAN E. MUSTAFA received the B.S. degree in electrical engineering from HITEC University Taxila, Pakistan, in 2012, and the Master of Science degree in electronic engineering from Ghulam Ishaq Khan (GIK) Institute of Engineering Sciences and Technology, Pakistan, in 2015, under the Partial Scholarship Scheme (PSS). He is currently pursuing the Ph.D. degree with the Pakistan Institute of Engineering and Applied Sciences (PIEAS). His research interests include fault diagnosis and tolerance control and alarm management. He also got a scholarship from Information Technology (IT) Endowment Fund during the Ph.D. degree. He was a recipient of travel grant from GIK Institute for presenting his research work at DICE Energy Exhibition and Competition, in 2014, organized by DICE Foundation, USA, and NED University, Pakistan.



ABDUL QAYYUM KHAN (Senior Member, IEEE) received the B.Sc. degree (Hons.) in electrical engineering from the University of Engineering and Technology (UET), Peshawar, Pakistan, in 2001, the M.Sc. degree in systems engineering from the Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan, in 2004, and the Ph.D. degree in electrical engineering from the University of Duisburg-Essen, Duisburg, Germany, in 2010. He is currently a Professor and the Head of the Department of Electrical Engineering, PIEAS. His research interests include fault diagnosis in technical processes, linear and nonlinear observer design, robust control of nonlinear systems, and LMI-based optimal design. He was the Chair of the Technical Committee of the IEEE International Conference on Recent Advances in Electrical Engineering, in 2015, 2017, 2018, and 2019, and the IEEE International Conference on Emerging Technologies, in 2016. He is the Chair of the IEEE Control System Society Karachi-Lahore-Islamabad Joint Chapter.

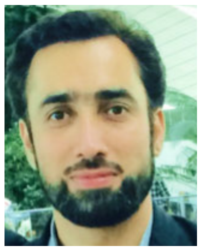


ABDUS SAMEE received the B.Sc. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, the master's degree in nuclear engineering from Quaid e Azam University Islamabad, and the Ph.D. degree in electrical power engineering from the Harbin Institute of Technology (HIT), China. His research interests include aging and degradation of HV insulation under non-uniform electric field, electrical systems modeling and simulation, distributed generation and integration, nuclear reactor operation and control, pulsed power, and reactor safety in the purview of severe accident management guidelines.

IJAZ AHMED received the B.S. degree in electrical engineering from the Federal Urdu University of Arts, Sciences and Technology (FUUAST), Islamabad, in 2010, and the M.S. degree in electrical engineering (power systems) from Bahria University, Islamabad, in 2014. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad. His research interests include consensus control of multiagent systems, artificial intelligence, group consensus, networked control systems, cyber security, modeling of sustainable power systems, machine learning, integration of renewable energy sources, energy storage and consumption, energy management, techno economics solutions, and green economy in the electricity market.

MUHAMMAD ABID (Member, IEEE) received the M.Sc. degree in systems engineering from the Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan, in 2004, and the Ph.D. degree in electrical engineering from the Institute of Automatic Control and Complex Systems, University of Duisburg-Essen, Duisburg, Germany, in 2010. He is currently a Professor with PIEAS. His research interests include model-based fault detection in nonlinear systems and robust and optimal control.

MOHAMMAD M. ALQAHTANI received the bachelor's degree in industrial engineering from King Khalid University, Abha, Saudi Arabia, in 2010, the M.S. degree in industrial engineering from The University of Arizona, Tucson, USA, in 2015, and the Ph.D. degree in industrial engineering from the University of Illinois at Chicago (UIC), in 2021. After that, he was appointed to the Department of Industrial Engineering, King Khalid University, in 2021, as an Assistant Professor, where he is currently a Professor in industrial engineering. He serves as a peer-reviewer for many scientific journals, including *Applied Energy* (Elsevier), *Journal of Cleaner Production* (Elsevier), *Intelligent Transportation Systems* (Wiley), *Wireless Networks* (Springer), and IEEE Transactions on Intelligent Transportation Systems.



MUHAMMAD KHALID (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the School of Electrical Engineering Telecommunications (EET), The University of New South Wales (UNSW), Sydney, Australia, in 2011. He was initially a Postdoctoral Research Fellow, for three years, and then he continued as a Senior Research Associate with the Australian Energy Research Institute, School of EET, UNSW, for two years. He is currently an Associate Professor with the Electrical Engineering Department, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia. He is also a Research Affiliate with the Interdisciplinary Research Center for Renewable Energy and Power Systems (IRC-REPS) and the SDAIA-KFUPM Joint Research Center for Artificial Intelligence, KFUPM. He has authored/coauthored several journals and conference papers in the field of control and optimization for renewable power systems. In addition, he has been a reviewer for numerous international journals and conferences. His research interests include the optimization and control of battery energy storage systems for large-scale grid-connected renewable power plants (particularly wind and solar), distributed power generation and dispatch, hybrid energy storage, hydrogen systems, EVs, AI, machine learning, and smart grids. He was a recipient of a highly competitive postdoctoral writing fellowship from UNSW, in 2010. He was a recipient of many academic awards and research fellowships.