IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 24 November 2022, accepted 13 December 2022, date of publication 15 December 2022,
date of current version 20 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3229617

RESEARCH ARTICLE

Interval-Valued SVM Based ABO for Fault Detection and Diagnosis of Wind Energy Conversion Systems

MAJDI MANSOURI[1,2], (Senior Member, IEEE), KHALED DHIBI[3], HAZEM NOUNOU[1], (Senior Member, IEEE), AND MOHAMED NOUNOU[4], (Senior Member, IEEE)

[1]Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar
[2]Department of Mathematical Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia
[3]Laboratory of Automatic Signal and Image Processing, National School of Engineers of Monastir, University of Monastir, Monastir 5019, Tunisia
[4]Chemical Engineering Program, Texas A&M University at Qatar, Doha, Qatar

Corresponding author: Majdi Mansouri (majdi.mansouri@qatar.tamu.edu)

This work was supported in part by the Qatar National Library, and in part by the Qatar National Research Fund (QNRF) Research Grant.

ABSTRACT In this paper, special attention is paid to the detection and diagnosis of various incipient faults of uncertain wind energy conversion (WEC) systems. The proposals will enhance the monitoring and diagnosis of the WEC system while taking into account the system uncertainties. The developed techniques are based on Support Vector Machine (SVM) model to improve the diagnosis of WEC systems. First, to deal with model uncertainties in WEC systems, the SVM will be extended to interval-valued data with the aim of achieving greater accuracy and robustness for these uncertainties. Then, to improve even more the performances of the developed interval-valued SVM, multiscale data representation will be used to develop multiscale extensions of interval-valued SVM. Next, as a feature selection tool, an improved extension of Artificial Butterfly Optimization (ABO) algorithm is used in order to extract the significant features from data and improve the diagnosis results of multiscale interval SVM. The proposed improved ABO method consists in reducing the number of samples in the training data set using the Euclidean distance and extracting the most significant features from the reduced data using ABO algorithm. This in turn plays a vital role in improving the accuracy using the multiscale interval-SVM method and reducing the computational and storage costs.

INDEX TERMS Support vector machine (SVM), fault diagnosis, fault classification, artificial butterfly optimization (ABO), wind energy conversion (WEC).

I. INTRODUCTION

The growing energy needs of modern society raise significant environmental concerns. Therefore, basic research on renewable energy technologies is important to realize the potential of cleaner energy resources. Over the past few decades, wind energy has become one of the most important types of renewable energy in the world [1]. Additionally, improvements in wind energy technologies have enabled Wind Energy Conversion (WEC) systems to play an important role in the global energy market [2]. At the heart of the WEC systems is the power converter, which is used as a link between the machine and the grid [3]. Due to its impact on power performance, the WEC Converter (WECC) should be reliable, and readily available. The authors in [4] reported that the majority of electrical failures of WECs occur in the semiconductor devices of both grid and generator converters. Therefore, the downtime becomes increasingly reliant on the WEC system size [5]. However, WEC systems suffer from repeated failure, due to undetectable periodic faults. The open-circuit, wear-out, and short-circuit are the three transistor faults considered in this paper.

In recent years, different Fault Detection and Diagnosis (FDD) algorithms were developed and implemented, but machine learning-based methods proved to be more efficacy [6], [7]. For instance, the authors in [8] proposed a new

The associate editor coordinating the review of this manuscript and approving it for publication was Mark Kok Yew Ng.

paradigm based on Recurrent Neural Networks (RNNs)) for diagnosing faults in WEC systems. The developed technique enhanced the application of RNN model by reducing the number of samples using Hierarchical K-means clustering. In one of these techniques, a hybrid approach that uses Gaussian Process Regression (GPR) and Random Forest (RF) methods for FDD of WEC systems is proposed [9]. In this proposal, GPR algorithm is used for feature extraction and selection and RF model is used for the classification tasks. A fault diagnosis technique based on Hidden Markov model (HMM) is proposed in [10]. In this work, Principal Component Analysis (PCA) is used as a feature extraction and selection tool in order to improve the classification accuracy of HMM model. PCA has been successfully used for dimensionality reduction. Despite the proven performance of the PCA technique, it is based on the principle that the system behaves linearly. However, most real process data requires nonlinear techniques. To address these issues, a FDD technique based on Kernel PCA (KPCA) method for features extraction and selection and RF for classification task is proposed in [11]. Despite the efficacy of the KPCA technique, it is very expensive in terms of computation time and storage costs which limits its effectiveness in real-world applications with massive data. Support Vector Machine (SVM) is used for fault Detection in Wind turbines in which two traditional and two new kernels are used [12]. In [13], fault detection and classification in wind turbine systems using a cuckoo-optimized SVM is developed. Recently different deep learning techniques like RNNs, and Convolutional Neural Network (CNN) are widely used in FDD of WEC systems [14], [15]. In this paper, we focus on the development of an improved SVM technique to diagnose WEC systems.

SVM is one of the most popular machines learning algorithms that is widely used in FDD of different industrial systems due to its high accuracy and robustness [16]. The main idea behind SVM algorithm is to perform a nonlinear classification using the kernel trick by mapping the data into high-dimensional feature spaces [16]. SVM is more efficacy in cases where the number of dimensions is greater than the number of samples. Additionally, SVM is relatively memory efficient when there is a clear margin of separation between classes. However, the SVM algorithm has some disadvantages. Firstly, the classical SVM-based FDD technique doesn't take into account the uncertainties of the system. Uncertainty in systems is the consideration of maximum and minimum values, while the representation of single-valued data is affected by simplifying the data exploration procedure. Then, SVM algorithm does not perform very well when the data set has more noise which limits the application of real-world applications. Finally, SVM is not suitable for large data sets. In addition, It will under-perform in the case when the number of features for each data point exceeds the number of training data samples. Therefore, a feature selection step that extracts a new subset of features from the original data with low dimensionality is important to address the last challenge.

The first objective of this work is to develop an improved extension of SVM classifier based on interval-valued data to deal with the problem of uncertainties in WEC systems. To do this, we develop a SVM model based on upper and lower bounds (SVM^{UL}) to highlight the use of interval-valued data instead of single-valued data. However, row data contain high levels of noise and autocorrelation that can significantly limit the efficacy of using of any classifier for classification. Therefore, in the second step, we propose to use a multiscale based Wavelet Transform (WT) representation. WT represents a powerful tool for transforming time-domain signals into the time-frequency domain to achieve efficient separation of deterministic features from random noise.

Recently, various optimization techniques have been proposed for feature selection to improve the effectiveness of machine learning methods for FDD [17], [18]. An optimization algorithm is a step that is executed iteratively by comparing different solutions till an optimum or favorable solution is found [18]. Optimization techniques (i.e Artificial Butterfly Optimization (ABO)) help machine learning algorithms to be more effective. ABO was developed first by [19]. It represents one of the most recently developed optimization algorithms which mimics the butterflies' behavior for food foraging [20]. It has been utilized for several problems and demonstrated its efficiency and performance compared to other optimization algorithms [19]. Different works have applied ABO for feature selection purposes. For example, the authors in [19] showed that ABO outperformed well-known optimization algorithms through different applications including classical engineering problems and benchmarks. In [21], the authors apply ABO to find the maximum power point tracking under partial shading condition in photovoltaic (PV) systems. Also in [22], the authors presented Automatic Generation Control to two nonlinear power systems using ABO. The authors in [23] applied an enhanced version by introducing the adaptive mechanism and applied it to some benchmark applications. In [24], a binary version of ABO is proposed for feature Selection (FS) problem. Besides, in [25], the authors developed an improved version of ABO using Cross-Entropy method to achieve a better balance between exploration and exploitation. In [26], the authors have applied ABO based method for fault-tolerant and accurate localization in mobility assisted underwater wireless sensor networks.

Therefore, to overcome the challenges when the number of features per data exceeded the number of training data samples and also to resolve the problem of SVM when dealing with large data sets, we propose a feature selection step using ABO algorithm. However, in high dimensional problems as a feature extractor, the ABO algorithm may be stuck in local optima. Also, it has a solutions diversity problem. Therefore, an improved optimization based ABO algorithm consists of reducing the original data set using the Euclidean Distance (ED) metric and selecting the most pertinent features from reduced data using the ABO algorithm. The main objective behind using the ED technique is to conserve only one

observation in case of redundancy. Therefore, the new reduced data set is characterized by more diversity between samples which makes the use of ABO for features selection more effective. Once the features are selected using ABO based algorithm, they are fed to SVM based interval-valued model for classification purposes. To summarise, this paper proposes an enhanced SVM using data pre-processing based on interval-valued representation and multiscale representation, and an improved ABO algorithm for features selection. Two kinds of classifiers are considered in this work: a multi-class classifier and a set of one-class classifiers. The multi-class classifier consists of classifying instances into one or more classes. To better improve the diagnosis abilities, a bank of one-class classifiers is proposed. The presented results prove that the proposed method offers enhanced diagnosis performances when applied to uncertain WEC systems.

The rest of the paper is structured as follows. Section 2 introduces a detailed description of the developed methods. In Section 3, the developed techniques are evaluated using the WEC system. Section 4 concludes the article.

II. PROPOSED METHODOLOGIES

This work focuses on the fault diagnosis of Wind Energy Conversion (WEC) systems under different operational conditions while taking into consideration system uncertainties. The developed algorithms will use Support Vector Machines (SVM) to improve the diagnosis of WEC systems. First, to deal with model uncertainties in WEC systems, the SVM model will be extended to interval-valued data with the aim of achieving greater robustness and accuracy. Next, to deal with the problems of high levels of noise and autocorrelation that mask the important features in the data, a multiscale representation tool is proposed to develop a multiscale extension of ISVM method. Wavelet Transformation (WT) represents a powerful tool to transform time-domain signals into a time-frequency domain in order to achieve efficient separation of the deterministic characteristics of random noise. Finally, to improve even more the performances of the proposed frameworks, a feature selection scheme using an improved Artificial Butterfly Optimization (ABO) Algorithm is proposed. The improved ABO algorithm combines the advantages of ABO method and Euclidean distance metric to enhance the performance of multiscale ISVM. The main objective behind the improved ABO algorithm for feature selection is to overcome the problem when using ISVM for large data sets and also when the number of features for each data point exceeds the number of training data samples. The improved ABO algorithm involves two main steps: i) reducing the number of samples in the training data set using the Euclidean distance method, this, in turn, plays a pivotal role in significantly reducing the size of data and increasing the diversity between samples, ii) extract the most significant features from the reduced data using ABO algorithm. The proposed technique is called reduced ABO-MS-ISVM.

A. INTERVAL SUPPORT VECTOR MACHINES (ISVM)

The interval representation of the data observation x_{ij} is is presented using the lower bound $\underline{x}_{ij}$ and the upper bound $\overline{x}_{ij}$ as follows,

$$[x_{ij}] = [\underline{x}_{ij}, \overline{x}_{ij}] \tag{1}$$

where $i = 1, .., N$ and $j = 1, \ldots, m$, is the $i-$th observation of the $j-$th variables.

For the interval X^{UL} technique, an upper-lower technique is presented to define the new data. The lower and upper bounds matrices X^L and X^U are computed as,

$$X^L = \begin{pmatrix} \underline{x}_{11} & . & . & \underline{x}_{1m} \\ . & . & & . \\ . & & . & . \\ \underline{x}_{N1} & . & . & \underline{x}_{Nm} \end{pmatrix} \tag{2}$$

$$X^U = \begin{pmatrix} \overline{x}_{11} & . & . & \overline{x}_{1m} \\ . & . & & . \\ . & & . & . \\ \overline{x}_{N1} & . & . & \overline{x}_{Nm} \end{pmatrix} \tag{3}$$

The interval matrix X^{UL} is constructed using the upper and lower matrices can at the same time as,

$$x_{ij}^{UL} = \theta \underline{x}_{ij} + (1-\theta)\overline{x}_{ij} \tag{4}$$

where, $\theta \in [0, 1]$, θ present the regulation weight of interval-valued data unit. The main objective of θ is to poise the relationship between the upper and lower bounds. The new constructed upper and lower matrix is presented as,

The upper and lower matrix is given by:

$$X^{UL} = \begin{pmatrix} x_{11}^{ul} & . & . & x_{1m}^{ul} \\ . & . & & . \\ . & & . & . \\ x_{N1}^{ul} & . & . & x_{Nm}^{ul} \end{pmatrix} \tag{5}$$

Consider a given training set of N samples $\{x_k^{ul}, x_k^{ul}\}_{k=1}^N$, with input data $x_k^{ul} \in \mathbb{R}^m$ and output $x_k^{ul} \in \{-1\ 1\}$ which represents a set of labeled training features. The ISVM takes the following form:

$$y_k^{ul} = f(x_k^{ul}) = w^T x_k^{ul} + b \tag{6}$$

where $w \in \mathbb{R}^{2m}$ and $b \in \mathbb{R}$. In order to orient the linear separation hyperplanes as far as possible from the support vectors, the margin defined by $\frac{1}{\|w\|}$ should be maximized (see Fig.1). It is equivalent to finding:

$$\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{subject to} \quad x_k^{ul}(w^T x_k^{ul} + b) - 1 \geq 0 \ \ \forall k \tag{7}$$

Allocating the constraints Lagrange multipliers α, where $\alpha_k \geq 0\ \forall k$, the quadratic programming (QP) optimization problem (7) becomes

$$L_{QP} = \frac{1}{2}\|w\|^2 - \sum_k \alpha_k y_k (w^T x_k^{ul} + b) + \sum_k \alpha_k \tag{8}$$

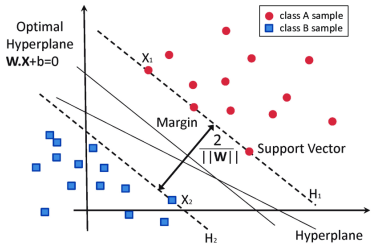


FIGURE 1. Support vector machine algorithm.

by setting the partials of (8), with respect to w and b, to zero

$$\begin{bmatrix} \frac{\partial L_{QP}}{\partial w} \\ \frac{\partial L_{QP}}{\partial b} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} w = \sum_k \alpha_k y_k^{ul} x_k^{ul} \\ \sum_k \alpha_k y_k^{ul} = 0 \end{cases} \tag{9}$$

Substituting (9) into (8) gives the dual optimization problem of the primary L_{QP} depending on α:

$$\max_{\alpha} \left[\sum_k \alpha_k - \tfrac{1}{2} \alpha^T H \alpha \right] \tag{10}$$

subject to

$$\sum_k \alpha_k y_k^{ul} = 0 \text{ and } \alpha \geq 0 \ \forall k$$

where $H = \left[y_k^{ul} y_t x_k^{ul} x_t \right]_{k=1,N}^{t=1,N}$. The solution of the optimization problem (10) returns to determine α. After that, w is easily calculated using (9). To determine b, a new data x_s satisfying (9) is presented, and a support vector takes the following form:

$$y_s(w^T x_s + b) = 1 \tag{11}$$

Referring to the separation hyperplane equations and using the class y_s, the variable b is given by:

$$b = y_s - \sum_{m \in S} \alpha_m y_m x_m x_s \tag{12}$$

In turn, the optimal orientation of the separating hyperplanes is obtained.

The main steps of the proposed interval SVM technique are summarized in the schematic diagram 2.

B. MULTISCALE INTERVAL SUPPORT VECTOR MACHINES (MS-ISVM)

In this step the interval-valued data is decomposed in multiple scales using wavelet based multiscale decomposition. By given a time domain data set (signal), the scaled signals can be obtained by projecting the original signal on a set of orthonormal scaling functions and convolution the original signal with a low pass filter (h) [27].

$$\phi_{i,j}(t) = \sqrt{2^{-j}} \phi(2^{-j} t - k) \tag{13}$$

where, k and j represent the dilation and translation parameters, respectively.

The difference between the original and approximated signals is called the detail signals. They are obtained by

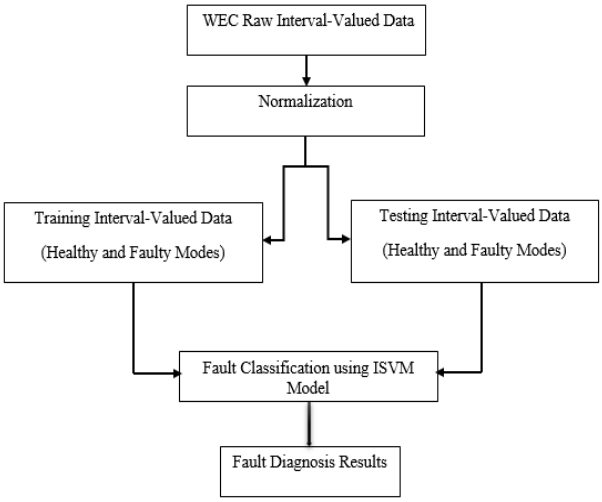


FIGURE 2. Schematic diagram of the ISVM technique.

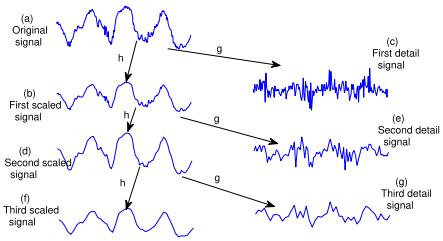


FIGURE 3. A schematic diagram of data representation at multiple scales [30].

projecting the original signal on a set of wavelet basis functions of the following equation [28]:

$$\psi_{i,j}(t) = \sqrt{2^{-j}} \psi(2^{-j} t - k) \tag{14}$$

Therefore, the original signal can be defined as the sum of the last scaled signal and all the detail signals of all scales [27]:

$$x(t) = \Sigma_{k=1}^{n2^{-J}} a_{jk} \phi_{jk} + \Sigma_{j=1}^{J} \Sigma_{k=1}^{n2^{-J}} d_{jk} \psi_{jk}(k) \tag{15}$$

where J and n are the number of scales and the length of the original signal, respectively [28], [29]. Figure 3 shows a schematic diagram that illustrates this multiscale representation process.

C. IMPROVED ARTIFICIAL BUTTERFLY OPTIMIZATION (ABO)

Generally, the lack of diversity between samples limits the effectiveness of ABO algorithm by causing the problem of local optimum areas due to premature convergence [31]. To overcome this challenge, we propose an improved ABO algorithm that involves two main steps. First, we reduce the number of observations from data using the Euclidean distance metric. This provides a significant impact to increase the diversity between samples and then avoid the problem

of local sub-optimal areas and premature convergence when using the ABO optimization algorithm.

Euclidean distance represents the shortest distance between two samples [32]. The dissimilarity matrix D that represent measurement of dissimilarity between all pairs of the samples for a data matrix X with N samples and m process variables is defined as,

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ . & . & . & . \\ d_{N1} & d_{N2} & \dots & d_{NN} \end{bmatrix} \tag{16}$$

where d_{ij} represents the ED between the rows X_i and X_j of the data matrix X. So, d_{ij} is computed as,

$$d_{ij} = \sqrt{\sum_{k=1}^{m} (X(i,k) - X(j,k))^2} \tag{17}$$

Thus, the new reduced data matrix X_r is presented as,

$$X_r = \left[x'(1)\ x'(2)\ \dots\ x'(N') \right]^T \in \mathbf{R}^{N' \times m} \tag{18}$$

where N' is the size of the new training data matrix. Then, using the ED metric all redundancies in samples are left out from the original data matrix. Next, we apply ABO algorithm to the reduced data. ABO is based on three assumptions: i) each butterfly moves randomly or towards the butterfly with the most scent, ii) all butterflies emit a scent and are attracted to each other, iii) A fitness function determines the intensity of the stimulus of the butterfly [33]. ABO solves the optimization problem through global and local search [34]. When a butterfly does not feel the fragrance of any butterfly it moves randomly to a new position in the search space, this procedure is called global search. When it senses the best butterfly's fragrance it moves toward it, this procedure is called local search.

The fragrance can be presented as follows,

$$g_i = cI^a, \qquad i = 1, \dots, N' \tag{19}$$

where N' is the number of butterflies, g_i represents the perceived magnitude of fragrance, c is the sensory modality, I represent the stimulus intensity, and a is the power exponent based on the degree of fragrance absorption.

The mathematical equation of the butterflies' global search movements is formulated as,

$$x_i^{t+1} = x_i^t + (r^2 * y_{best} - x_i^t) * g_i \tag{20}$$

where x_i^t represents the solution vector x_i of the ith butterfly in t iteration, r is a random number equal 0 or 1, y_{best} represents the current best solution, and g_i represent the fragrance of the ith butterfly. The mathematical equation of the butterflies' local search movements is formulated as,

$$x_i^{t+1} = x_i^t + (r^2 * x_i^k - x_j^t) * g_i \tag{21}$$

where x_i^k and x_j^t are ith and jth butterflies chosen randomly. The butterfly becomes a local random walk when x_i^k and x_j^t belong to the same iteration. In the case when x_i^k and x_j^t do not belong to the same iteration, the nature of random movement will diversify the solution.

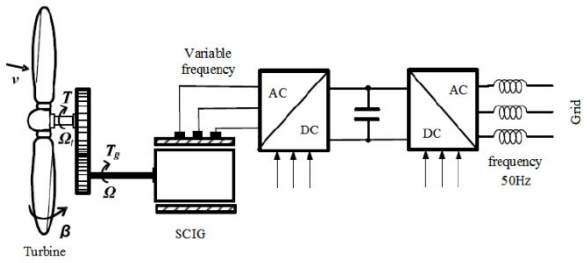


FIGURE 4. Diagram of variable speed wind turbine based on asynchronous machine.

TABLE 1. Main electrical faults in wind energy converters.

Fault symbol	description ($IGBT_{11}$)	Fault symbol	description ($IGBT_{21}$)
SC_{11}	Short-Circuit	SC_{21}	Short-Circuit
OC_{11}	Open-Circuit	OC_{21}	Open-Circuit
WO_{11}	Wear-Out	WO_{21}	Wear-Out

III. EXPERIMENTAL RESULTS AND COMPARATIVE STUDY

A. SYSTEM DESCRIPTION

In this proposal, special attention is paid to diagnosing various incipient faults during different modes of operation in WEC systems.

The open-circuit, wear-out, and short-circuit are the three transistor faults considered in this paper.

A description of the variable speed wind turbine based on asynchronous machine is shown in Figure 4.

The model of the system contains two parts. One part consists of the model of the turbine and the Squirrel Cage Induction Machine (SCIG). This part of the system will be controlled through the stator side AC/DC Converter. The second part of the model consists of the grid side DC/AC converter sub-system. This configuration allows unlimited variable speed operation. The generated voltage is rectified and transformed into direct current and voltage whatever the rotation speed of the machine. The detailed description of the turbine is given in [10]. In the wind chain, the power converters topology is on two levels (Figure 5). Each converter consists of three arms. Each arm contains a high and a low IGBTs. In this study, the diagnosis study concerns the IGBT11 for the generator converter and the IGBT21 for the grid converter. In this case, we considered three types of faults (short-circuit, open-circuit and wear-out) and their description is provided in the Table 1. The last fault is modeled by an internal resistance equal to two Ohms. The behavior of some electrical and mechanical variables for different fault scenarios are is presented in Figures 6 to 10.

In this study, 12 variables have been generated for modeling and fault classification as listed in Table 2 [35]. To demonstrate the effectiveness of the developed methodologies, real bearing vibration data are used as an example [10].

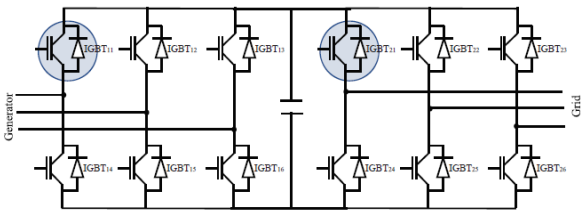


FIGURE 5. Converters topology.

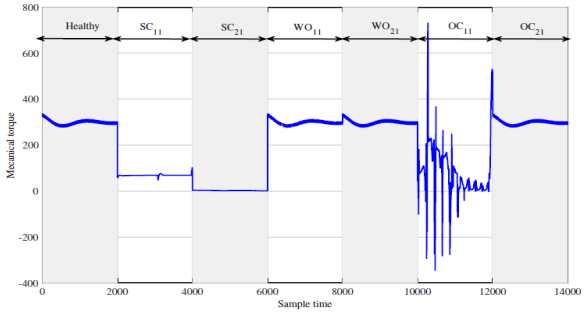


FIGURE 6. Mechanical torque under dissimilar conditions [10].

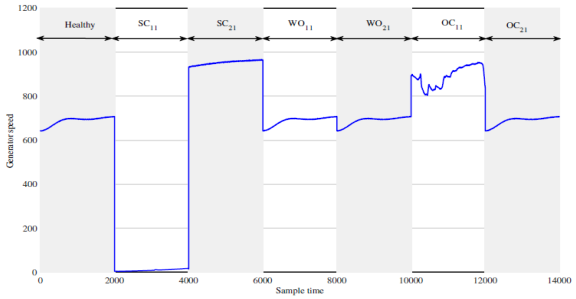


FIGURE 7. Generator speed under dissimilar conditions [10].

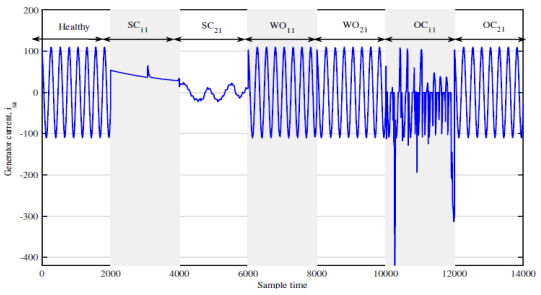


FIGURE 8. Generator current under dissimilar conditions [10].

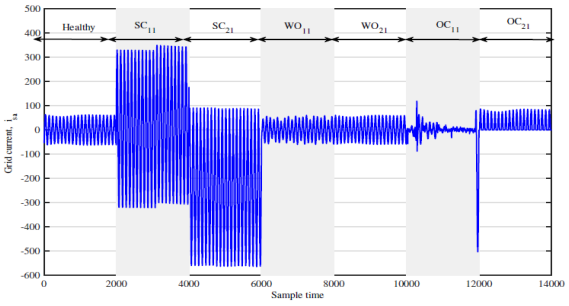


FIGURE 9. Grid current under dissimilar conditions [10].

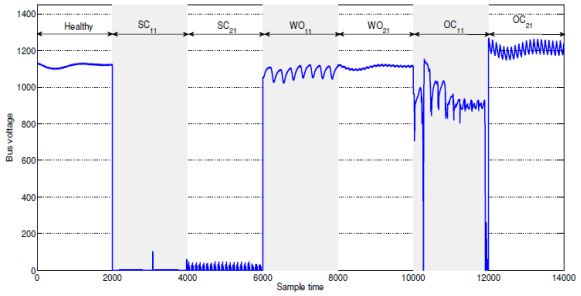


FIGURE 10. Bus voltage under dissimilar conditions [10].

TABLE 2. Variables description.

Variables	description
C_m	Mechanical torque (Nm)
N_g	Generator speed (tr/m)
i_{sag}	Generator current phase a (A)
i_{sbg}	Generator current phase b (A)
I_{sd}	Generator current along d-axis (A)
I_{sq}	Generator current along q-axis (A)
V_{DC}	Bus voltage (V)
P_{Out}	Output power (W)
i_{sar}	Grid current phase a (A)
i_{sbr}	Grid current phase b (A)
I_{sd}	Grid current along d-axis (A)
I_{sq}	Grid current along q-axis (A)

TABLE 3. Construction of database.

State	Process status	Training	Testing
C_0	Healthy	2000	2000
C_1	Faulty-SC_{11}	2000	2000
C_2	Faulty-SC_{21}	2000	2000
C_3	Faulty-WO_{11}	2000	2000
C_4	Faulty-WO_{21}	2000	2000
C_5	Faulty-OC_{11}	2000	2000
C_6	Faulty-OC_{21}	2000	2000

B. RESULTS AND DISCUSSIONS

In this work, 7 operating modes including one healthy and 6 separate faulty operating modes of the WEC system are tested in order to make simulation data series (Table 3). Each operating mode is adequately qualified over 2000 10-time-lagged observations during a second time period with 20 KHz as sampling frequency.

To evaluate the performances of the proposed techniques in terms of accuracy, recall, and precision a 10-fold cross-validation criterion was adopted. In this study, for the SVM model, the Kernel is selected using the radial basis function, and the K and C parameters for SVM are chosen with the lowest RMSE value. One healthy case (assigned to class C0) and six different faulty cases (assigned to C_1 up to C_6) are considered (Table 3). The number of hidden layers chosen

for artificial neural network techniques and the RNN model is ten and the number of hidden neurons in the hidden layer is equal to 50. To investigate our methods performance in different multi-classes classification and compare that with different existing techniques, classification results in terms of accuracy, recall, and precision and also in terms of computation time are summarized in Table 4. From Table 4, it can be seen that the proposed interval SVM provides the best results in terms of accuracy (94.25%), recall (94.25%), and precision (94.42%) compared to the SVM for single-valued data which gives values of accuracy, recall, and precision equal to 92.14%, 92.14%, and 92.16%, respectively. The presented results given by the proposed ISVM demonstrate the impact of using interval-valued data instead of single-valued data in improving the classification results. To show the impact of using multiscale representation, the presented results given by MS-ISVM are compared with those given by ISVM. The presented results from Table 4 demonstrate that the developed MS-ISVM (96.60%/96.60%/96.63%) provides an important enhancement in terms of accuracy, recall, and precision compared to ISVM method. These results demonstrate the effectiveness of multiscale and interval-valued techniques to enhance the performance of SVM model. To further illustrate the efficacy of the model, ABO-MS-ISVM technique based on using of ABO algorithm for feature selection, multiscale representation, and interval-valued data was compared with MS-ISVM and ISVM to demonstrate the effectiveness of ABO algorithm to enhance the fault diagnosis abilities using MS-ISVM model. The presented results in Table 4 show that the proposed ABO-MS-ISVM performed with better accuracy (98.85%), recall(98.85%), and precision (98.91%) compared to MS-ISVM. As can be seen, the diagnosis results of the proposed RABO-MS-ISVM are 99.34%, 99.34%, and 99.35% in terms of accuracy, recall, and precision, respectively. The presented results given by RABO-MS-ISVM technique demonstrate the effectiveness of using the improved ABO algorithm for feature selection that is based on the ED method as a reduction tool in improving the diagnosis results. In addition, it is clear from Table 4 that the proposed RABO-MS-ISVM and ABO-MS-ISVM techniques outperform other existing techniques for example RNN. The poor classification results using SVM model are due to the direct use of measured variables which indicates the success of the developed methods which select the more pertinent features before performing the classification task using SVM model. On the other hand, the presented results demonstrate that the developed RABO-MS-ISVM provides an important reduction in terms of computation time (5.47s) compared to ABO-MS-ISVM technique (12.39s). Thus, the proposed RABO-MS-ISVM not only reduced the computation time but also provide a little improvement in the classification results compared to ABO-MS-ISVM method. To further intuitively reflect the fault diagnosis ability of the proposed techniques the confusion matrix results are presented in Tables 5, 6, 7, and 8. The confusion matrix presents the correctly classified observations and misclassified

TABLE 4. Performances comparison of different multi-class techniques.

Method	Accuracy	Recall	Precision	CT(s)
RABO-MS-ISVM	99.34	99.34	99.35	5.47
ABO-MS-ISVM	98.85	98.85	98.91	12.39
MS-ISVM	96.60	96.60	96.63	15.27
ISVM	94.25	94.25	94.42	16.04
SVM	92.14	92.14	92.16	15.69
KNN	88.30	88.31	88.31	0.91
NN	93.70	93.70	93.73	3.53
FFNN	97.17	97.17	97.8	8.14
CFNN	97.17	97.17	97.16	8.36
GRNN	97.06	97.06	97.06	7.06
RNN	98.15	98.16	98.15	18.93

TABLE 5. Confusion matrix of RABO-MS-ISVM.

Conf. Matrix		Predicted class							Recall
True classes	C0	1992	0	0	0	8	0	0	99.60
	C1	0	2000	0	0	0	0	0	100
	C2	0	0	2000	0	0	0	0	100
	C3	20	0	0	1932	48	0	0	96.60
	C4	0	0	0	0	2000	0	0	100
	C5	0	0	0	16	0	1984	0	99.20
	C6	0	0	0	0	0	0	2000	100
Precision		99.01	100	100	99.18	97.28	100	100	**99.34**

observations for healthy mode (C_0) and faulty modes (C_1 to C_6). Table 5 which presents the confusion matrix of the developed RABO-MS-ISVM shows that for the healthy case (C_0) the recall is 99.60% and the precision is 99.01% with 0.40% of misclassification. In this case, 1992 observations are identified as healthy among 2000 samples. In this table, it can be seen for the faulty case C_5, the recall is 99.20%, the precision is 100%, 1984 samples are identified as fault C_5 among 2000 and 16 samples are identified as fault C_3. Using the ABO-MS-ISVM technique, as seen in Table 6 during the healthy case (C_0), 1926 samples are identified as healthy among 2000 samples and 74 samples are classified as C_4. Table 7 shows that during the healthy case (C_0), MS-ISVM method identify 1857 samples among 2000 and 27 samples are identify as fault C_3 and 114 samples are identify as fault C_4. As shown in Table 8, ISVM method identify only 1752 samples from 2000 for healthy case (C_0) with a misclassification equal to 12.40%. As shown in Tables 5, 6, 7, and 8, the proposed ABO-MS-ISVM and RABO-MS-ISVM techniques present perfect results through all used classification metrics for all operating modes.

To further highlight the diagnosis effectiveness of the proposed techniques, a one-class classifier framework is presented. Table 9 presents the global performance accuracy using four proposed techniques in the case of one class

TABLE 6. Confusion matrix of ABO-MS-ISVM.

Conf. Matrix		Predicted class							Recall
True classes	C0	1926	0	0	0	74	0	0	96.30
	C1	0	2000	0	0	0	0	0	100
	C2	0	0	2000	0	0	0	0	100
	C3	10	0	0	1927	63	0	0	96.35
	C4	13	0	0	0	1987	0	0	99.35
	C5	0	0	0	0	0	2000	0	100
	C6	0	0	0	0	0	0	2000	100
Precision		98.82	100	100	100	93.55	100	100	**98.85**

TABLE 7. Confusion matrix of MS-ISVM.

Conf. Matrix		Predicted class							Recall
True classes	C0	1857	0	0	29	114	0	0	92.85
	C1	0	2000	0	0	0	0	0	100
	C2	0	0	2000	0	0	0	0	100
	C3	5	0	0	1884	111	0	0	94.20
	C4	63	0	0	153	1784	0	0	89.20
	C5	0	0	0	0	0	2000	0	100
	C6	0	0	0	0	0	0	2000	100
Precision		96.47	100	100	91.19	88.80	100	100	**96.60**

TABLE 8. Confusion matrix of ISVM.

Conf. Matrix		Predicted class							Recall
True classes	C0	1752	0	0	28	216	0	4	87.60
	C1	4	1996	0	0	0	0	0	99.80
	C2	4	0	1996	0	0	0	0	88.20
	C3	168	0	4	1764	60	4	0	86.00
	C4	224	0	4	48	1720	0	4	99.60
	C5	4	4	0	0	0	1992	0	98.80
	C6	8	0	0	12	0	4	1976	100
Precision		80.96	99.80	99.60	95.25	86.17	99.60	99.60	**94.25**

TABLE 9. Accuracy with different one class classifiers.

	Methods			
Class	RABO-MS-ISVM	ABO-MS-ISVM	MS-ISVM	ISVM
C0	99.68	97.59	93.61	88.12
C1	99.97	99.95	99.98	99.87
C2	99.99	99.97	99.99	90.07
C3	97.11	97.43	97.05	94.17
C4	99.91	99.42	89.46	89.86
C5	99.27	99.97	99.98	99.91
C6	99.98	99.97	99.98	99.97
Mean accuracy	99.41	99.18	97.15	94.56
Mean CT	4.08	10.27	14.03	15.14

classifier scenario. As presented in Table 9, the proposed ABO-MS-ISVM and RABO-MS-ISVM methods achieve a mean of accuracy equal to 99.41% and 99.18%, respectively. Besides, we can conclude from Table 9, that the proposed paradigms, which are based on the feature selection step using ABO algorithm, outperform the other two diagnostic techniques (MS-ISVM and ISVM) in terms of mean of classification accuracy. Moreover, we can show that the proposed RABO-MS-ISVM based on data reduction framework provides an important reduction in terms of computation time (4.08s) compared to ABO-MS-ISVM (10.27s) technique. We can conclude from the results presented in Table 9, that the proposed RABO-MS-ISVM present the best tread-off between high diagnosis metrics and low computation time. It is due to the efficacy and the impact of using the improved ABO algorithm for features selection, multiscale representation, and interval-valued techniques to afford the best tread-off between high diagnosis metrics and low computation time.

IV. CONCLUSION

This work proposed a new approaches aimed at improving the diagnosis of WEC systems by developing intelligent and innovative WEC fault diagnosis frameworks. For this purposes, an enhanced SVM classifier using metaheuristic optimization, data reduction, multiscale data representation, and interval-valued representation is proposed. First, SVM-based interval upper and lower bounds technique is developed to deal with model uncertainties in WEC systems. Then, a multiscale extension of ISVM method is developed in order to deal with the problems of high levels of noise and autocorrelation that mask the important features in the data and improve the classification results. Finally, a feature selection tool using an improved ABO algorithm is developed. The improved ABO algorithm for feature selection consists of reducing the data using the Euclidean distance metric in order to maximize the diversity between data samples. Then ABO method is used to select the most pertinent features from reduced data to perform the classification task. The proposed methodologies not only improved the diagnosis abilities but also reduced the computation time and storage cost. The experimental results demonstrated the feasibility and effectiveness of the developed frameworks.

REFERENCES

[1] J. Brożyna, W. Strielkowski, A. Fomina, and N. Nikitina, "Renewable energy and EU 2020 target for energy efficiency in the Czech republic and Slovakia," *Energies*, vol. 13, no. 4, p. 965, Feb. 2020.

[2] L. Zhang, H. Zhang, and G. Cai, "The multiclass fault diagnosis of wind turbine bearing based on multisource signal fusion and deep learning generative model," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[3] D. Zappalá, N. Sarma, S. Djurović, J. C. Crabtree, A. Mohammad, and P. J. Tavner, "Electrical & mechanical diagnostic indicators of wind turbine induction generator rotor faults," *Renew. Energy*, vol. 131, pp. 14–24, Feb. 2019.

[4] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis—Part I: Components and subsystems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6536–6545, Oct. 2015.

[5] J. Lan, R. J. Patton, and X. Zhu, "Fault-tolerant wind turbine pitch control using adaptive sliding mode estimation," *Renew. Energy*, vol. 116, pp. 219–231, Feb. 2018.

[6] T. Sun, G. Yu, M. Gao, L. Zhao, C. Bai, and W. Yang, "Fault diagnosis methods based on machine learning and its applications for wind turbines: A review," *IEEE Access*, vol. 9, pp. 147481–147511, 2021.

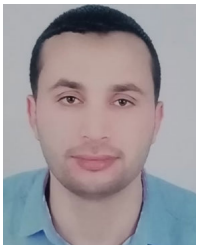
[7] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, and G. Nenadic, "Machine learning methods for wind turbine condition monitoring: A review," *Renew. Energy*, vol. 133, pp. 620–635, Apr. 2019.

[8] M. Mansouri, K. Dhibi, M. Hajji, K. Bouzara, H. Nounou, and M. Nounou, "Interval-valued reduced RNN for fault detection and diagnosis for wind energy conversion systems," *IEEE Sensors J.*, vol. 22, no. 13, pp. 13581–13588, Jul. 2022.

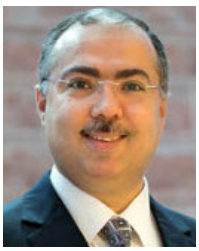
[9] M. Mansouri, R. Fezai, M. Trabelsi, H. Nounou, M. Nounou, and K. Bouzrara, "Reduced Gaussian process regression based random forest approach for fault diagnosis of wind energy conversion systems," *IET Renew. Power Gener.*, vol. 15, no. 15, pp. 3612–3621, Nov. 2021.

[10] A. Kouadri, M. Hajji, M.-F. Harkat, K. Abodayeh, M. Mansouri, H. Nounou, and M. Nounou, "Hidden Markov model based principal component analysis for intelligent fault diagnosis of wind energy converter systems," *Renew. Energy*, vol. 150, pp. 598–606, May 2020.
[11] R. Fezai, K. Dhibi, M. Mansouri, M. Trabelsi, M. Hajji, K. Bouzrara, H. Nounou, and M. Nounou, "Effective random forest-based fault detection and diagnosis for wind energy conversion systems," *IEEE Sensors J.*, vol. 21, no. 5, pp. 6914–6921, Mar. 2021.
[12] P. Santos, L. F. Villa, A. Reñones, A. Bustillo, and J. Maudes, "An SVM-based solution for fault detection in wind turbines," *Sensors*, vol. 15, no. 3, pp. 5627–5648, 2015.
[13] A. Agasthian, R. Pamula, and L. A. Kumaraswamidhas, "Fault classification and detection in wind turbine using Cuckoo-optimized support vector machine," *Neural Comput. Appl.*, vol. 31, no. 5, pp. 1503–1511, May 2019.
[14] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3196–3207, Apr. 2019.
[15] J. Lei, C. Liu, and D. Jiang, "Fault diagnosis of wind turbine based on long short-term memory networks," *Renew. Energy*, vol. 133, pp. 422–432, Apr. 2018.
[16] M. Pal and P. M. Mather, "Support vector machines for classification in remote sensing," *Int. J. Remote Sens.*, vol. 26, no. 5, pp. 1007–1011, 2005.
[17] B. Zheng, H.-Z. Huang, W. Guo, Y.-F. Li, and J. Mi, "Fault diagnosis method based on supervised particle swarm optimization classification algorithm," *Intell. Data Anal.*, vol. 22, no. 1, pp. 191–210, Feb. 2018.
[18] W. Tuerxun, X. Chang, G. Hongyu, J. Zhijie, and Z. Huajian, "Fault diagnosis of wind turbines based on a support vector machine optimized by the sparrow search algorithm," *IEEE Access*, vol. 9, pp. 69307–69315, 2021.
[19] M. Mafarja, I. Aljarah, A. A. Heidari, H. Faris, P. Fournier-Viger, X. Li, and S. Mirjalili, "Binary dragonfly optimization for feature selection using time-varying transfer functions," *Knowl.-Based Syst.*, vol. 161, pp. 185–204, Dec. 2018.
[20] S. Arora and S. Singh, "Butterfly optimization algorithm: A novel approach for global optimization," *Soft Comput.*, vol. 23, no. 3, pp. 715–734, 2018.
[21] K. Aygül, M. Cikan, T. Demirdelen, and M. Tumay, "Butterfly optimization algorithm based maximum power point tracking of photovoltaic systems under partial shading condition," *Energy Sources A, Recovery, Utilization, Environ. Effects*, pp. 1–19, 2019, doi: 10.1080/15567036.2019.1677818.
[22] D. K. Lal, A. Barisal, and S. D. Madasu, "AGC of a two area nonlinear power system using BOA optimized FOPID+PI multistage controller," in *Proc. 2nd Int. Conf. Adv. Comput. Commun. Paradigms (ICACCP)*, Feb. 2019, pp. 1–6.
[23] B. Singh and P. Anand, "A novel adaptive butterfly optimization algorithm," *Int. J. Comput. Mater. Sci. Eng.*, vol. 7, no. 4, Dec. 2018, Art. no. 1850026.
[24] S. Arora and P. Anand, "Binary butterfly optimization approaches for feature selection," *Expert Syst. Appl.*, vol. 116, pp. 147–160, Feb. 2019.
[25] G. Li, F. Shuang, P. Zhao, and C. Le, "An improved butterfly optimization algorithm for engineering design problems using the cross-entropy method," *Symmetry*, vol. 11, no. 8, p. 1049, Aug. 2019.
[26] M. Nain, N. Goyal, S. Rani, R. Popli, I. Kansal, and P. Kaur, "Hybrid optimization for fault-tolerant and accurate localization in mobility assisted underwater wireless sensor networks," *Int. J. Commun. Syst.*, vol. 35, no. 17, Nov. 2022, Art. no. e5320.
[27] M. N. Nounou, H. N. Nounou, N. Meskin, A. Datta, and E. R. Dougherty, "Multiscale denoising of biological data: A comparative analysis," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 9, no. 5, pp. 1539–1545, Sep. 2012.
[28] M. N. Nounou and H. N. Nounou, "Reduced noise effect in nonlinear model estimation using multiscale representation," *Model. Simul. Eng.*, vol. 2010, pp. 1–8, Jan. 2010.
[29] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, no. 7, pp. 909–996, 1988.
[30] M. Z. Sheriff, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Fault detection using multiscale PCA-based moving window GLRT," *J. Process Control*, vol. 54, pp. 47–64, Jun. 2017.
[31] A. S. Assiri, "On the performance improvement of butterfly optimization approaches for global optimization and feature selection," *PLoS ONE*, vol. 16, no. 1, Jan. 2021, Art. no. e0242612.
[32] K. Elmore and M. Richman, "Euclidean distance as a similarity metric for principal component analysis," *Monthly Weather Rev.*, vol. 129, no. 3, pp. 540–549, 2001.
[33] X. Qi, Y. Zhu, and H. Zhang, "A new meta-heuristic butterfly-inspired algorithm," *J. Comput. Sci.*, vol. 23, pp. 226–239, Nov. 2017.
[34] A. A. Awad, A. F. Ali, and T. Gaber, "Feature selection method based on chaotic maps and butterfly optimization algorithm," in *Proc. Int. Conf. Artif. Intell. Comput. Vis.* Cham, Switzerland: Springer, 2020, pp. 159–169.
[35] K. Dhibi, M. Mansouri, K. Bouzrara, H. Nounou, and M. Nounou, "Reduced neural network based ensemble approach for fault detection and diagnosis of wind energy converter systems," *Renew. Energy*, vol. 194, pp. 778–787, Jul. 2022.

MAJDI MANSOURI (Senior Member, IEEE) received the degree in electrical engineering from SUPCOM, Tunis, Tunisia, in 2006, the M.Sc. degree in electrical engineering from ENSEIRB, Bordeaux, France, in 2008, the Ph.D. degree in electrical engineering from UTT Troyes, France, in 2011, and the H.D.R. (Accreditation To Supervise Research) degree in electrical engineering from the University of Orleans, France, in 2019. He joined the Electrical Engineering Program, Texas A&M University at Qatar, in 2011, where he is currently an Associate Research Scientist. He is the author of more than 200 publications. He is also the author of the book *Data-Driven and Model-Based Methods for Fault Detection and Diagnosis* (Elsevier, 2020). His research interests include development of model-based, data-driven, and machine learning techniques for fault detection and diagnosis.



KHALED DHIBI received the Ph.D. degree in electronics and microelectronics from the Faculty of Sciences of Monastir (FSM), University of Monastir, Monastir, Tunisia. His work focuses on the implementation of data-driven techniques for fault detection and diagnosis of industrial processes.



HAZEM NOUNOU (Senior Member, IEEE) is currently a Professor in electrical and computer engineering at Texas AM University at Qatar. He has more than 19 years of academic and industrial experience. He has served as an associate editor and on the technical committees of several international journals and conferences. He has significant experience in research on control systems, databased control, system identification and estimation, fault detection, and system biology. He has been awarded several NPRP research projects in these areas. He has successfully served as the lead PI and a PI on five QNRF projects, some of which were in collaboration with other PI's in this proposal. He has published more than 200 refereed journals and conference papers and book chapters.

MOHAMED NOUNOU (Senior Member, IEEE) is currently a Professor in chemical engineering at Texas A&M University at Qatar (TAMU). He has more than 19 years of combined academic and industrial experience. His research interests include the area of systems engineering and control, with emphasis on process modeling, monitoring, and estimation. He has published more than 200 refereed journals and conference publications and book chapters. He has successfully served as the lead PI and a PI on several QNRF projects (six NPRP projects and three UREP projects). He is a Senior Member of the American Institute of Chemical Engineers (AIChE).