

Received 31 January 2023, accepted 10 February 2023, date of publication 14 February 2023, date of current version 21 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3244838

## RESEARCH ARTICLE

# Enhanced PSO-Based NN for Failures Detection in Uncertain Wind Energy Systems

KHALED DHIBI<sup>1</sup>, MAJDI MANSOURI<sup>1</sup>, (Senior Member, IEEE),  
MOHAMED TRABELSI<sup>2</sup>, (Senior Member, IEEE), KAMALELDIN ABODAYEH<sup>3</sup>,  
KAIS BOUZRARA<sup>4</sup>, HAZEM NOUNOU<sup>1</sup>, (Senior Member, IEEE),  
AND MOHAMED NOUNOU<sup>5</sup>, (Senior Member, IEEE)

<sup>1</sup>Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar

<sup>2</sup>Department of Electronics and Communications Engineering, Kuwait College of Science and Technology, Kuwait City 27235, Kuwait

<sup>3</sup>Department of Mathematical Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia

<sup>4</sup>Research Laboratory of Automation, Signal Processing and Image, National Engineering School of Monastir, Monastir 5019, Tunisia

<sup>5</sup>Chemical Engineering Program, Texas A&M University at Qatar, Doha, Qatar

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

This work was supported by the Qatar National Library through the Qatar National Research Fund (QNRF) Research Grant.

**ABSTRACT** Ensuring the validity of measurements in wind energy systems (WES) is a challenging task in system diagnosis and data validation. This work, therefore, elaborates on the development of new approaches aimed at improving the operation of WES by developing intelligent and innovative fault diagnosis frameworks. Therefore, an enhanced particle swarm optimization (PSO), data reduction, and interval-valued representation are proposed. First, a feature selection tool using PSO Algorithm is developed. Then, in order to maximize the diversity between data samples and improve the effectiveness of using PSO algorithm for feature selection, the Euclidean distance metric is used in order to reduce the data and maximize the diversity between data samples. Finally, PSO and RPSO-based interval centers and ranges and upper and lower bounds techniques are developed to deal with model uncertainties in WES. The last retained features from the proposed PSO-based methods are fed to the neural network (NN) classifier. The proposed methodology improves the diagnosis abilities, reduces the computation time, and decreased the storage cost. The presented experimental results prove the high performance of the suggested paradigms in terms of computation time and accuracy.

**INDEX TERMS** Uncertainties, interval-valued data, wind energy systems, particle swarm optimization (PSO), neural network (NN), feature selection, dataset size reduction.

## LIST OF ABBREVIATIONS AND ACRONYMS

FD	Fault Detection.
WES	Wind Energy Systems.
ML	Machine Learning.
IVD	Interval-Valued Data.
PSO	Particle Swarm Optimization.
RPSO	Reduced PSO.
IPSO	Interval PSO.
IRPSO	Interval Reduced PSO.
NN	Neural Network.

ANN	Artificial NN.
ED	Euclidean Distance.
CR	Centers and Ranges.
UL	Upper and Lower.
CT	Computation Time.
CM	Confusion Matrix.

## I. INTRODUCTION

Nowadays, Wind Energy systems (WES) have been attracting an increased interest with the aim of reducing carbon dioxide emissions [1], [2], [3], [4]. However, the effective operations of WES are facing several challenges mainly related to wind intermittency, vibrations, and unexpected failures in power converters [5], [6]. Thus, effective fault detection (FD)

The associate editor coordinating the review of this manuscript and approving it for publication was Ton Duc Do<sup>1</sup>.

techniques should be applied [7]. Generally, two main classes of FD techniques can be identified: model-based and data-driven methods. Model-based approaches are based on a consistency test between the observed process behavior of the sensor and the expected behavior of a mathematical process model, which is generally derived using a basic understanding (using physical and chemical principles) of the process under fault-free conditions. However, these methods belong to the capability of the mathematical model to correctly characterize the system's behaviors. Data-driven techniques are based on historical data collected during fault-free process operations [8]. Using the training data, an empirical model is built and used for faults detection by exploiting the future measurement data [8].

The most well-known data-driven diagnosis techniques applied to detect faults in WES include machine learning and statistical process control [9]. For instance, density-based spatial clustering was used to classify faults state, and decision tree (DT) and random forest (RF) methods were used to build the predictive models for WES anomalies. In other works, a diagnostic technique based on the signal analysis method has been developed [10]. This technique eliminates redundant information using an empirical mode decomposition and discrete wavelet transform (DWT) technique. In [11], a fault diagnosis technique includes support vector machine (SVM) model and the optimal composition of symptom parameters algorithm to diagnose the motor fault in WES. However, this proposal suffers from some limitations due to the selection of the features which may lead to the misclassification of faults. To overcome the limits of using classical SVM, a fault diagnosis technique that proposes a large margin distributed machine (LDM) and introduces margin variance and mean on the basis of SVM is presented [12]. This proposal shows the best classification results compared to classical SVM. A decision tree has been developed to diagnose faults in WT systems [13]. Although this method is easy to perform, it has drawbacks in dealing with missing values. Another fault detection method based on artificial neural networks (ANN) is presented in [14]. Recently, deep learning techniques such as convolution neural network (CNN), recurrent neural network, and long short-term memory (LSTM) have been widely used in fault diagnosis of WES [15], [16], [17]. In [18], a fusion neural network model is developed by combining long short-term memory neural network (LSTM NN) and broad learning system (BLS) algorithm, to outstanding predict the lithium-ion battery capacity and remaining useful life (RUL). In [19], wavelet packet decomposition is used to extract characteristics that are fed to CNN classifier to perform the classification task. Deep learning techniques can accurately detect several conventional faults, but they have some limitations represented in the hard training step and the high time complexity. Different ways may improve the use of machine learning (ML) for FD purposes. Generally, the existing intelligent FD methods consist of two main steps: feature selection and fault classification. In [5],

a fault diagnosis technique for WES based on an ML algorithm is developed to solve performance problems flexibly and reliably. In this case, the reduced kernel principal components algorithm is used to extract the most significant features from raw data and perform the classification task using a random forest classifier. This proposal presents high diagnosis accuracy but it needs a high computation time. Another fault diagnosis methodology based on features extraction and selection step is introduced in [20]. This technique consists of extracting and selecting features from raw sensor data using an improved Gaussian process regression (GPR). The main disadvantage of Gaussian processes is the loss of efficiency in high-dimensional spaces. In [21], a hybrid approach is proposed by combining the variational modal decomposition (VMD), particle filter (PF), and GPR to forecasting battery future capacity and remaining useful life (RUL). In our previous work [22], an efficient feature selection technique based on particle swarm optimization (PSO) is proposed. The basic objective behind the PSO algorithm is to remove irrelevant features and extract only the most significant ones from raw data in order to improve the classification task using a neural network classifier. Besides, in this proposal, an improved extension of the PSO model based on the use of Euclidean distance (ED) is developed. The main idea behind the use of the ED method is to avoid the problem of premature convergence and local sub-optimal areas when using the classical PSO optimization algorithm. Also, the proposed reduced PSO-NN (RPSO-NN) algorithm aims to improve the results in terms of accuracy as well as in terms of complexity time, and storage cost. Generally, uncertainty and inaccuracy might describe the significant information characterizing the real systems [23]. Classical data is a simplification during the data mining process and it may cause severe loss of information. For this reason, interval-valued data representation is important [24], [25]. Recently, different faults diagnosis techniques have been proposed to deal with noise in data, imbalanced data types and inadequate fault sample data [26], [27]. In [26], an improved label-noise robust generative adversarial network was used to ensure the quality of the generated data and improve the generalizability of the model under actual operating conditions scenarios by performing a batch comparison between generated and actual data.

In this paper, different classification paradigms, merging the benefits of Particle Swarm Optimization (PSO) and Neural Networks (NN) algorithms, are proposed for classification of faults in uncertain WES. The first developed method, so-called PSO-based NN, is addressed so that PSO is applied for feature selection, and then the selected pertinent features are fed to the NN classifier to perform the classification task. Unfortunately, the application of PSO-based NN algorithms in complex problems is limited by the lack of diversity causing premature convergences (local sub-optimal areas). To overcome this limitation, an Euclidean distance (ED)-based data reduction tool is used.

One of the main objectives of this work is to improve data-based diagnosis techniques in uncertain WES. Generally, uncertainty and inaccuracy might describe the significant information characterizing the real systems [23]. Classical data is a simplification during the data mining procedure and it may cause severe loss of information. Thus, interval-valued data representation is of high importance [24], [25]. In this work, more robust techniques can be achieved by describing the process measurements by interval-valued data instead of single-valued data. Therefore, interval models are to be applied to handle the new nature of data.

The objective is to improve data-based monitoring methods, especially in experimental industrial applications with imperfect measurements which significantly worsen the fault diagnosis task. The developed techniques are able to improve both fault diagnosis robustness and sensitivity while maintaining a satisfactory and stable performance over long periods of process operation. To summarize, the main paper's contributions are as follows:

- 1) An effective PSO-based feature selection method is proposed. The aim is to remove irrelevant features and only extract the most significant ones from raw data (better classification accuracy using NNs).
- 2) An Euclidean distance (ED)-based reduced PSO-NN (RPSO-NN) algorithm is proposed (higher accuracy, lower complexity, and reduced storage cost) to overcome the premature convergence and local sub-optimum limitations of classical PSO optimization algorithms.
- 3) The proposed techniques are extended to interval-valued data to deal with model uncertainties such as noise, measurement errors, and variability.

The proposed techniques are characterized by their high performance in terms of accuracy, robustness, and ability to detect incipient/drift faults and their severity in WES.

This paper is organized as follows: Section III presents the proposed techniques while the description of the studied WES is presented in Section III. Then, Section IV presents the results and the performance evaluation. Finally, Section V concludes the paper.

## II. INTERVAL ENHANCED PSO-BASED NN FOR FAULT DETECTION

### A. PROPOSED DIAGNOSIS PARADIGM

In this paper, the aim is to detect incipient faults while taking into consideration the WES uncertainties. The developed interval enhanced PSO based NN technique exploits the benefits of features selection based on PSO, dataset size reduction using ED, interval-valued representation of raw data, and NN for classification. Firstly, the proposed algorithms will use PSO method for features selection and neural networks classifier for classification to improve the diagnosis of WES. PSO is a very effective global search technique based on the movement and intelligence of swarms (number of samples). An Euclidean distance (ED)-based reduced PSO-NN (RPSO-NN) algorithm is proposed (higher accuracy, lower

### Algorithm 1 IRPSO-Based NN Algorithm

Input  $N \times m$  data matrix  $X$ .

1. Normalize the data set,
2. Determine the new interval data matrix  $(X^{UL}, X^{CR})$ ,
3. Compute the reduced matrix using the euclidean distance metric by conserving only one observation in the case of redundancy,
4. Select features using PSO model,
5. Use the selected features as input to the NN classifier for the training phase,
6. The NN classifier is evaluated using testing features,
7. Classifying the different operating modes (healthy and faulty).

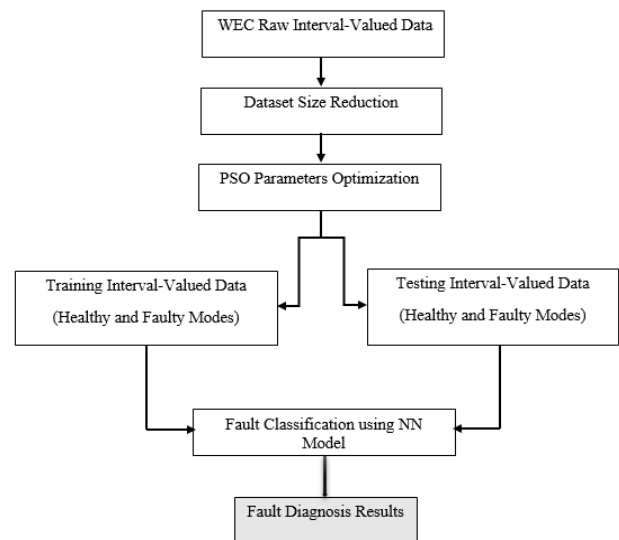


FIGURE 1. Schematic diagram of the interval reduced PSO-NN techniques.

complexity, and reduced storage cost) to overcome the premature convergence and local sub-optimum limitations of classical PSO optimization algorithms. The proposed technique is extended to interval-valued data to deal with model uncertainties such as noise, measurement errors, and variability. The developed approaches for interval-valued data are IPSO-NN<sub>CR</sub>, IPSO-NN<sub>UL</sub>, IRPSO-NN<sub>CR</sub>, and IRPSO-NN<sub>UL</sub>. The selected features are then fed to the NN classifier.

The main steps of the proposed IRPSO-NN technique are summarized in algorithm 1 and schematic diagram 1.

### B. TECHNIQUES

In this section, a brief description of the proposed methods is presented.

#### 1) ENHANCED PARTICLE SWARM OPTIMIZATION (PSO)

Particle Swarm Optimization (PSO) is considered as a stochastic optimization technique based on the intelligence and movement of swarms [28], [29]. PSO is a very efficient global search algorithm that needs very few algorithm

parameters compared to other optimization algorithms like genetic algorithms which require setting different evolutionary operators such as crossover and mutation [29].

To overcome the limitation of premature convergence (due to the lack of diversity) which limits the effectiveness of PSO algorithms, an ED-based data reduction tool is proposed. The main objective of the ED method is to extract only a single sample in case of redundancy to construct the reduced data. This, in turn, plays a pivotal role to increase the diversity between data samples and then solving the problem of premature convergence and local sub-optimal area. Let's consider a data matrix  $X$  with  $N$  samples and  $m$  process variables, the ED between the rows  $X_i$  and  $X_j$  of the data matrix  $X$  is computed by:

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

The dissimilarity matrix  $D$  representing the dissimilarity between all samples' pairs for a data matrix  $X$  is computed by:

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ \dots & \dots & \dots & \dots \\ d_{N1} & d_{N2} & \dots & p_{NN} \end{bmatrix} \quad (2)$$

Thus, we obtain a reduced data matrix  $X'$  with  $N'$  samples and  $m$  process variables where  $N' < N$ . Then we apply the PSO algorithm for selecting the more pertinent features from the reduced data.

## 2) INTERVAL REDUCED PSO (IRPSO)

The uncertainties in the collected data present the interval-valued data (IVD). IVD methodology is very important to preserve variable information. In this step, four techniques IPSO<sub>CR</sub>, IPSO<sub>UL</sub>, IRPSO<sub>CR</sub>, and IRPSO<sub>UL</sub> are used. The main idea behind these proposals is to use a specific interval-valued data matrix instead of the single-valued data matrix. The interval data matrices are  $X_{CR}$  and  $X_{UL}$ .  $X_{CR}$  constructed by the concatenation of center and range matrices and  $X_{UL}$  constructed by the lower and upper bounds of interval values of variables.

Let's consider  $x_{ij}$ , where  $i = 1, \dots, N$  and  $j = 1, \dots, m$ , is the  $i$ -th sample of the  $j$ -th observation, the interval representation of the data observation  $x_{ij}$  is defined using the lower bound  $\underline{x}_{ij}$  and the upper bound  $\bar{x}_{ij}$  as follows,

$$[x_{ij}] = [\underline{x}_{ij}, \bar{x}_{ij}] \quad (3)$$

The interval data matrix  $[X]$  is constructed using the interval-valued samples for the different description variables as follows:

$$[X] = \begin{pmatrix} [\underline{x}_{11}, \bar{x}_{11}] & \dots & [\underline{x}_{1m}, \bar{x}_{1m}] \\ \vdots & \ddots & \vdots \\ [\underline{x}_{N1}, \bar{x}_{N1}] & \dots & [\underline{x}_{Nm}, \bar{x}_{Nm}] \end{pmatrix} \\ = ([x_1], \dots, [x_N])^T \quad (4)$$

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,

$$X^L = \begin{pmatrix} \underline{x}_{11} & \dots & \underline{x}_{1m} \\ \vdots & \ddots & \vdots \\ \underline{x}_{N1} & \dots & \underline{x}_{Nm} \end{pmatrix} \quad (5)$$

$$X^U = \begin{pmatrix} \bar{x}_{11} & \dots & \bar{x}_{1m} \\ \vdots & \ddots & \vdots \\ \bar{x}_{N1} & \dots & \bar{x}_{Nm} \end{pmatrix} \quad (6)$$

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

$$x_{ij}^{UL} = \theta \underline{x}_{ij} + (1 - \theta) \bar{x}_{ij} \quad (7)$$

where,  $\theta \in [0, 1]$ ,  $\theta$  represents the regulation weight of interval-valued data unit. The new constructed upper and lower matrix is computed by:

$$X^{UL} = \begin{pmatrix} x_{11}^{UL} & \dots & x_{1m}^{UL} \\ \vdots & \ddots & \vdots \\ x_{N1}^{UL} & \dots & x_{Nm}^{UL} \end{pmatrix} \quad (8)$$

It is worth mentioning that  $\theta = 1$  represents a lower scheme with one feature while  $\theta = 0$ , represents an upper bound including the size information of  $\bar{x}$ .

The interval  $[x_{jk}]$  can be also represented as a couple  $\{x_{jk}^c, x_{jk}^r\}$ . The center  $x_{jk}^c$  and the range  $x_{jk}^r(k)$  of the interval are computed respectively as,

$$x_j^c(k) = \frac{1}{2}(\bar{x}_{jk} + \underline{x}_{jk}) \quad (9)$$

$$x_j^r = \frac{1}{2}(\bar{x}_{jk} - \underline{x}_{jk}) \quad (10)$$

The center and range matrices from interval-valued data matrix are computed as,

$$X^c = \frac{1}{2} \begin{pmatrix} \underline{x}_{11} + \bar{x}_{11} & \dots & \underline{x}_{1m} + \bar{x}_{1m} \\ \vdots & \ddots & \vdots \\ \underline{x}_{N1} + \bar{x}_{N1} & \dots & \underline{x}_{Nm} + \bar{x}_{Nm} \end{pmatrix} \quad (11)$$

$$X^r = \frac{1}{2} \begin{pmatrix} \bar{x}_{11} - \underline{x}_{11} & \dots & \bar{x}_{1m} - \underline{x}_{1m} \\ \vdots & \ddots & \vdots \\ \bar{x}_{N1} - \underline{x}_{N1} & \dots & \bar{x}_{Nm} - \underline{x}_{Nm} \end{pmatrix} \quad (12)$$

Then, we concatenate the computed center and range matrices into one matrix  $X^{CR}$  as,

$$X^{CR} = [X^c \ X^r] \in \mathbf{R}^{N \times 2m} \quad (13)$$

## 3) ARTIFICIAL NEURAL NETWORK

Neural network techniques have been widely used for FD in various applications [30], [31]. Neural networks (NNs) contain different layers, including an input layer, one or more hidden layers, and an output layer [31]. Neurons and

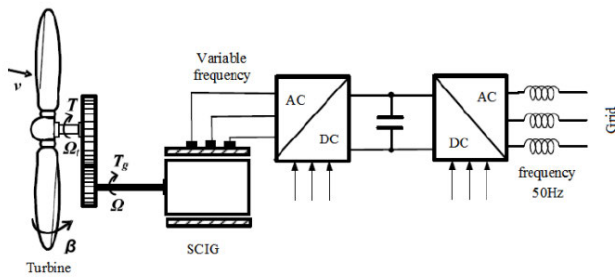


FIGURE 2. Synoptic of the WES under study.

weighted connections between neurons are the main components of neural networks [31]. The weights and the input-output function are the main evaluation criteria of the NN performance [31]. The main known Network architectures are feedforward and recurrent architecture. The main difference between the two architectures types present in the feedback between networks which is present when using recursive architecture by taking the correct prediction during backpropagation. In this paper, multilayer artificial neural networks (ANN) with Levenberg-Marquardt Backpropagation (LMBP) training method is adopted [32]. The developed ANN is constructed with 10 hidden layers and 50 hidden neurons in the hidden layer.

### III. SYSTEM DESCRIPTION

The WES under study is presented in Figure 2, while Figure 3 illustrates the topology of the used back-to-back converter. The detailed description of the WES under study was presented in [1].

WTC includes two main parts. The first one is the model of the turbine and the squirrel cage induction machine (SCIG). In this case, the stator side AC/DC Converter is used for the control. The second one is 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. A detailed description of the turbine is presented in [1]. In the wind chain, the power converters topology is on two levels (Figure 3). Each converter consists of three arms. Each arm contains high and low IGBTs. Failures in WEC systems are mainly caused by variations in weather conditions and faults in power converters. Recent studies have shown that more than 21% of the failures in WEC systems are caused by faults in the power conversion stage [33], though the WT and the gearbox failures cause the longest downtime [34]. Many elements might cause the fatigue of the switching devices. The fatigue impacts in particular the dynamics of the component and consequently may result in extra switching losses and even system failure. The fatigue is usually modeled by increasing the internal resistance of the component. In this work, it is assumed that the internal resistance is equal to zero during normal operating conditions. Then, the resistance increases to reflect fatigue.

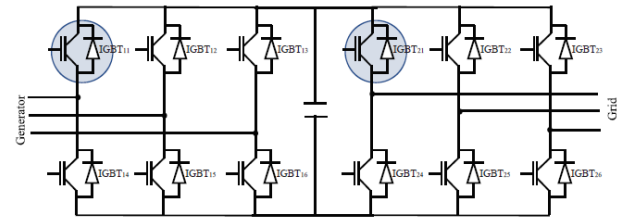


FIGURE 3. Back-to-back converter topology.

TABLE 1. Emulated faults.

Fault symbol	description ( <i>IGBT<sub>11</sub></i> )	Fault symbol	description ( <i>IGBT<sub>21</sub></i> )
<i>SC<sub>11</sub></i>	Short-Circuit	<i>SC<sub>21</sub></i>	Short-Circuit
<i>OC<sub>11</sub></i>	Open-Circuit	<i>OC<sub>21</sub></i>	Open-Circuit
<i>WO<sub>11</sub></i>	Wear-Out	<i>WO<sub>21</sub></i>	Wear-Out

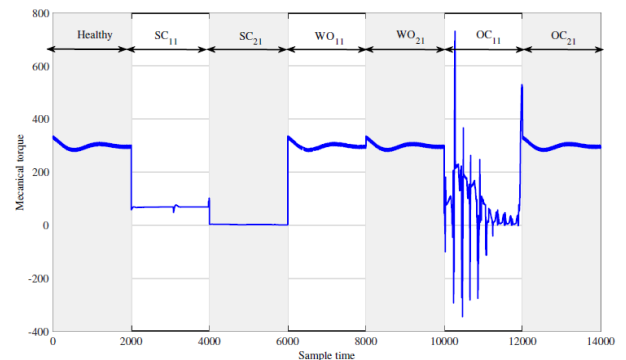


FIGURE 4. Mechanical torque under dissimilar conditions [1].

Therefore, it is necessary to detect the onset of fatigue to prevent the overall failure of the converter. Several electrical and mechanical variables need to be analyzed in detail to narrow down the failed parts. For instance, the collector-emitter voltage of an IGBT rises sharply just before the failure occurs. This can be used as an excellent indicator for preventive maintenance purposes in wind systems [35]. For the sake of simplicity, only IGBT11 (rectifier stage) and IGBT21 (inverter stage) are encompassed in the FD study. Three types of faults are considered (open-circuit, wear-out, and short-circuit) (Table 1). It is worth noting that the wear-out fault is modeled by an internal resistance of 2 Ω.

The behavior of some electrical and mechanical variables for different fault scenarios are presented in Figures 4 to 8.

In this study, 12 variables have been generated for modeling and fault classification as listed in Table 2 [36]. To demonstrate the effectiveness of the developed methodologies, real bearing vibration data are used as an example [1].

### IV. RESULTS AND DISCUSSIONS

Seven working modes (1 healthy and 6 faulty modes) are evaluated as illustrated in Table 3. Each operating mode is adequately qualified over 2000 10-time-lagged observations



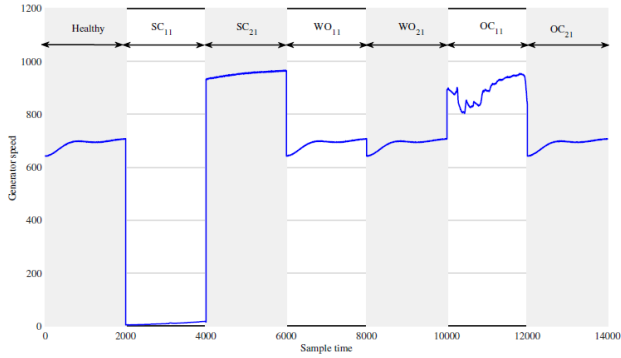


FIGURE 5. Generator speed under dissimilar conditions [1].

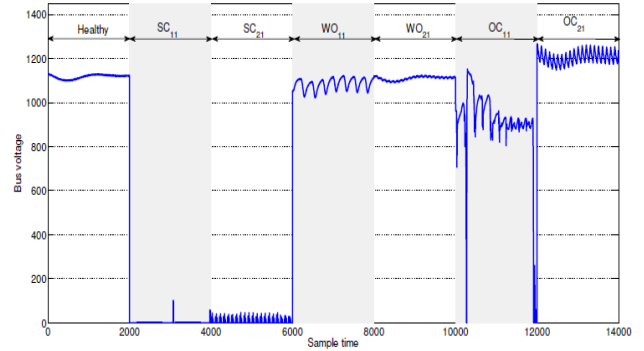


FIGURE 8. Bus voltage under dissimilar conditions [1].

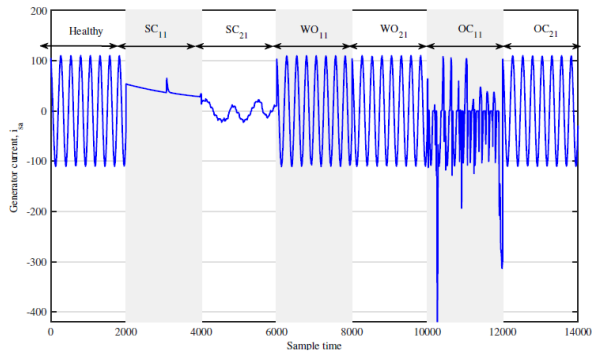


FIGURE 6. Generator current under dissimilar conditions [1].

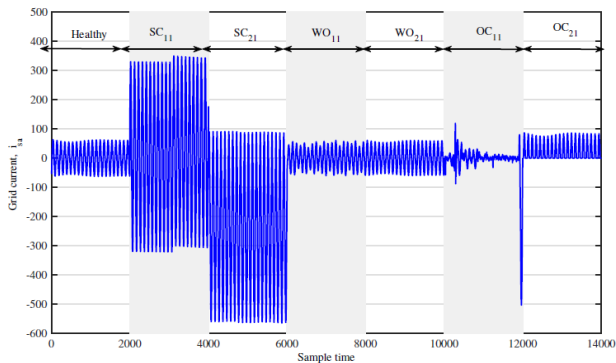


FIGURE 7. Grid current under dissimilar conditions [1].

during a second time period with 20 KHz as sampling frequency.

The swarm size, max iteration, cognition coefficient, and social coefficient of the PSO algorithm are equal to 20, 100, 2, and 2, respectively. The number of hidden layers of the NN structure was set to 10 while the number of hidden neurons is equal to 50 in the hidden layer.

In the multi-class classification stage, one healthy case (C0) and 6 faulty cases (C1-C6) are considered (Table 3). Table 4 shows that the proposed algorithms outperform the NN algorithm in terms of accuracy in the training and testing phases. Besides, The PSO-based NN provides the best results

TABLE 2. Variables description.

Variables	description
$C_m$	Mechanical torque (Nm)
$N_g$	Generator speed ( $tr/m$ )
$i_{saq}$	Generator current phase a (A)
$i_{sbq}$	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

compared to other optimization techniques like genetic algorithm (GA), differential evolution (DE), ant Colony optimization (ACO), simulated annealing (SA) and gravitational search algorithm (GSA) based methods. In addition, the presented results demonstrate that the proposed reduced PSO-based NN (RPSO-based NN) (19.87/1.19) provides an important reduction in terms of computation time compared to PSO-based NN (36.14/2.47) technique.

To prove the effectiveness of using interval-valued data compared to single-valued data, the results of the proposed interval techniques are compared to the proposed single-valued techniques. One can notice from Table 4 that both the proposed techniques for interval-valued data IPSO-based  $NN_{CR}$  (98.59/98.50) and IPSO-based  $NN_{UL}$  (98.73/98.75) provide the best accuracy results compared to PSO-based NN (98.20/98.16). Besides, a significant reduction in computation time is obtained by both IRPSO-based  $NN_{CR}$  (21.67/1.28) and IRPSO-based  $NN_{UL}$  (23.17/1.35)

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

Method	Training		Testing	
	Accuracy	CT(s)	Accuracy	CT(s)
IRPSO-based NN <sub>CR</sub>	99.18	21.67	99.15	1.28
IRPSO-based NN <sub>UL</sub>	99.71	23.17	99.68	1.35
IPSO-based NN <sub>CR</sub>	98.59	38.86	98.50	2.68
IPSO-based NN <sub>UL</sub>	98.73	40.03	98.75	2.76
RPSO-based NN	98.47	19.87	98.52	1.19
PSO-based NN	98.20	36.14	98.16	2.47
NN	93.19	45.11	93.70	3.53
GA-based NN	97.07	38.36	97.02	3.91
DE-based NN	96.82	36.84	96.83	2.87
ACO-based NN	96.88	37.14	96.84	2.53
SA-based NN	94.01	32.16	93.97	2.12
GSA-based NN	94.06	34.84	94.11	2.52
RNN	97.70	369.12	98.15	28.11
GRNN	96.79	99.14	97.01	6.97
SVM	83.74	128.75	92.14	15.69
RF	84.65	111.08	93.15	13.76
KNN	77.99	6.04	88.30	0.91

compared to IPSO-based NN<sub>CR</sub> (38.86/2.68) and IPSO-based NN<sub>UL</sub> (40.03/2.76) techniques for training and testing phases. In addition, The proposed IRPSO-based NN techniques not only decrease the computation time but also enhance the accuracy. The obtained accuracy using both IRPSO-based NN<sub>CR</sub> and IRPSO-based NN<sub>UL</sub> is equal to 99.18/99.15 and 99.71/99.68 for training and testing phases, respectively. A comparison study between the proposed methodologies and other existing methods such as recurrent neural network (RNN), General regression neural network (GRNN), SVM, RF, and KNN is presented. KNN and SVM present a low accuracy and they are not able to differentiate between the different operating modes. RNN classifier presents good results in terms of accuracy, but it is suffer from a high computation time for both the training and testing phases.

It is clear from Table 4 that the proposed IRPSO-based NN<sub>UL</sub> method achieves the best accuracy compared to other techniques.

In the second stage a one-class classification is done (Table 5). In this case, each classifier is trained to classify a specific class with a label of 1 or -1. Table 6 summarizes the results in terms of accuracy and mean of computation time of the developed techniques. The presented results from Table 6 prove that all the proposed techniques provide good results in terms of accuracy. In addition, we can conclude from the presented results in Table 6, that the developed techniques based on data-size reduction tool afford a good reduction in terms of computation time and with almost the same accuracy. Thus, the developed techniques based

**TABLE 5. Multiple one class classifier logic for fault diagnosis.**

Classifier for	Classes						
	C0	C1	C2	C3	C4	C5	C6
C0	-1	1	1	1	1	1	1
C1	1	-1	1	1	1	1	1
C2	1	1	-1	1	1	1	1
C3	1	1	1	-1	1	1	1
C4	1	1	1	1	-1	1	1
C5	1	1	1	1	1	-1	1
C6	1	1	1	1	1	1	-1

**TABLE 6. Accuracy with different one class techniques.**

Class	Phase	Methods					
		PSO-based NN	RPSO-based NN	IPSO-based NN <sub>CR</sub>	IPSO-based NN <sub>UL</sub>	IRPSO-based NN <sub>CR</sub>	IRPSO-based NN <sub>UL</sub>
C <sub>0</sub>	Training	98.23	98.49	98.65	98.71	99.22	99.83
	Testing	98.27	98.57	98.50	98.74	99.20	99.72
C <sub>1</sub>	Training	98.17	98.42	98.53	98.74	99.10	99.61
	Testing	98.14	98.17	98.47	98.62	98.99	99.64
C <sub>2</sub>	Training	97.01	97.57	98.18	97.43	98.07	98.28
	Testing	96.73	96.83	96.82	97.15	98.58	98.76
C <sub>3</sub>	Training	98.28	98.56	98.59	98.80	99.28	99.91
	Testing	98.34	98.52	98.57	98.74	99.09	99.90
C <sub>4</sub>	Training	98.23	98.53	98.71	98.69	99.31	99.90
	Testing	98.29	98.53	98.68	98.79	99.39	99.94
C <sub>5</sub>	Training	98.17	98.28	98.57	98.71	99.09	99.80
	Testing	98.30	98.01	98.29	98.67	99.15	99.86
C <sub>6</sub>	Training	98.27	98.49	98.76	98.85	99.36	99.94
	Testing	98.36	98.61	98.57	98.83	99.43	99.97
Mean accuracy	Training	98.05	98.33	98.57	98.56	99.06	99.61
	Testing	98.06	98.17	98.27	98.50	99.61	99.68
Mean CT (s)	Training	24.29	14.07	26.34	28.90	16.49	17.03
	Testing	1.23	0.67	1.47	1.60	0.74	0.82

on PSO algorithm for features selection and ED for data-size reduction can strongly decrease the size of the dataset while keeping the more informative features. In addition, the presented results demonstrate that the proposed IRPSO-based NN<sub>UL</sub> technique gives the best tread off between accuracy (99.71/99.68) and computation time (23.17/1.35) for training and testing phases.

**V. CONCLUSION**

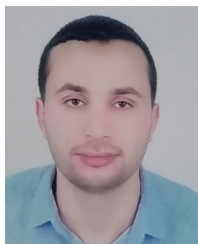
Diagnosis in WES is important to ensure reliable power production and optimal energy harvesting because WES usually suffer from several faults due to a difficult outdoor environment. This work elaborates on the development of new approaches aimed at improving the operation of WES by developing intelligent and innovative WES fault diagnosis frameworks. To do this, an enhanced NN-based classifier using PSO, and data reduction are proposed. The improved PSO algorithm is applied to enhance the classification accuracy by removing the irrelevant features and extracting the most significant ones from reduced data using the Euclidean distance method. The reduced PSO-NN (RPSO-NN) technique is characterized by higher accuracy and reduced

computation time and storage cost. To deal with model uncertainties (measurement errors, noise, variable variability, ...) in WES, the developed technique will be extended to interval-valued data with the aim of achieving greater accuracy. Therefore, based on the application of the interval-valued dataset, four techniques were proposed to deal with uncertainties in WES. The proposed methodologies not only improve the diagnosis abilities but also reduce the computation time and storage cost. The presented results demonstrate the effectiveness of the proposed approaches for fault diagnosis of WES. As future works, we will explore other machine learning and deep learning algorithms through single and interval-valued representation, allowing the comparison between the fault diagnosis performances of each algorithm for certain and uncertain WES systems. Finally, the developed fault diagnosis techniques will be utilized in practice to help improve operations of WEC systems. They will be tested and validated using simulated and real data under extreme conditions and using different simulation scenarios.

## REFERENCES

- [1] 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.
- [2] M. Tang, Q. Zhao, H. Wu, Z. Wang, C. Meng, and Y. Wang, "Review and perspectives of machine learning methods for wind turbine fault diagnosis," *Frontiers Energy Res.*, vol. 9, p. 596, Nov. 2021.
- [3] G. Chen and X. Cai, "Reconfigurable control for fault-tolerant of parallel converters in PMSG wind energy conversion system," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 604–614, Apr. 2019.
- [4] C. Xiao, Z. Liu, T. Zhang, and X. Zhang, "Deep learning method for fault detection of wind turbine converter," *Appl. Sci.*, vol. 11, no. 3, p. 1280, Jan. 2021.
- [5] R. Fezai, K. Dhibi, M. Mansouri, M. Trabelsi, M. Hajji, K. Bouzara, 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.
- [6] M. Mansouri, R. Fezai, M. Trabelsi, M. Hajji, M.-F. Harkat, H. Nounou, M. N. Nounou, and K. Bouzara, "A novel fault diagnosis of uncertain systems based on interval Gaussian process regression: Application to wind energy conversion systems," *IEEE Access*, vol. 8, pp. 219672–219679, 2020.
- [7] S. Chaudhary, R. Teodorescu, P. Rodriguez, P. C. Kjaer, and A. M. Gole, "Negative sequence current control in wind power plants with VSC-HVDC connection," *IEEE Trans. Sustain. Energy*, vol. 3, no. 3, pp. 535–544, Jul. 2012.
- [8] V. Venkatasubramanian, R. Rengaswamy, S. N. Kavuri, and K. Yin, "A review of process fault detection and diagnosis: Part III: Process history based methods," *Comput. Chem. Eng.*, vol. 27, no. 3, pp. 327–346, 2003.
- [9] J.-Y. Hsu, Y.-F. Wang, K.-C. Lin, M.-Y. Chen, and J. H.-Y. Hsu, "Wind turbine fault diagnosis and predictive maintenance through statistical process control and machine learning," *IEEE Access*, vol. 8, pp. 23427–23439, 2020.
- [10] A. Ouanas, A. Medoued, M. Mordjaoui, A. Lebaroud, and D. Sayad, "Fault diagnosis in yaw drive induction motor for wind turbine," *Wind Eng.*, vol. 42, no. 6, pp. 576–595, Dec. 2018.
- [11] H. T. Xue, M. Wang, Z. X. Li, and P. Chen, "Sequential fault detection for sealed deep groove ball bearings of in-wheel motor in variable operating conditions," *J. Vibroeng.*, vol. 19, pp. 5947–5959, Apr. 2017.
- [12] M. Tang, Z. Kuang, Q. Zhao, H. Wu, and X. Yang, "Fault detection of wind turbine pitch system based on multiclass optimal margin distribution machine," *Math. Problems Eng.*, vol. 2020, pp. 1–10, Aug. 2020.
- [13] I. Abdallah, V. Nertimanis, C. Mylonas, K. Tatsis, E. Chatzi, N. Dervilis, W. Keith, and M. Eoghan, "Fault diagnosis of wind turbine structures using decision tree learning algorithms with big data," in *Safety and Reliability—Safe Societies in a Changing World*. Boca Raton, FL, USA: CRC Press, 2018, pp. 3053–3061.
- [14] A. G. Kavaz and B. Barutcu, "Fault detection of wind turbine sensors using artificial neural networks," *J. Sensors*, vol. 2018, pp. 1–11, Dec. 2018.
- [15] S. M. Mousavi, M. Ghasemi, M. D. Manshadi, and A. Mosavi, "Deep learning for wave energy converter modeling using long short-term memory," *Mathematics*, vol. 9, no. 8, p. 871, Apr. 2021.
- [16] C. Gu and H. Li, "Review on deep learning research and applications in wind and wave energy," *Energies*, vol. 15, no. 4, p. 1510, Feb. 2022.
- [17] 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.
- [18] S. Zhao, C. Zhang, and Y. Wang, "Lithium-ion battery capacity and remaining useful life prediction using board learning system and long short-term memory neural network," *J. Energy Storage*, vol. 52, Aug. 2022, Art. no. 104901.
- [19] G. Q. 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. 2018.
- [20] A. Eftekhari, R. A. Hauser, and A. Grammenos, "MOSES: A streaming algorithm for linear dimensionality reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 11, pp. 2901–2911, Nov. 2020.
- [21] C. Zhang, S. Zhao, and Y. He, "An integrated method of the future capacity and RUL prediction for lithium-ion battery pack," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 2601–2613, Mar. 2022.
- [22] M. Mansouri, K. Dhibi, H. Nounou, and M. Nounou, "An effective fault diagnosis technique for wind energy conversion systems based on an improved particle swarm optimization," *Sustainability*, vol. 14, no. 18, p. 11195, Sep. 2022.
- [23] A. Emami-Naeini, M. M. Akhter, and S. M. Rock, "Effect of model uncertainty on failure detection: The threshold selector," *IEEE Trans. Autom. Control*, vol. AC-33, no. 12, pp. 1106–1115, Dec. 1988.
- [24] K. Dhibi, R. Fezai, M. Mansouri, A. Kouadri, M.-F. Harkat, K. Bouzara, H. Nounou, and M. Nounou, "A hybrid approach for process monitoring: Improving data-driven methodologies with dataset size reduction and interval-valued representation," *IEEE Sensors J.*, vol. 20, no. 17, pp. 10228–10239, Sep. 2020.
- [25] M.-F. Harkat, M. Mansouri, M. Nounou, and H. Nounou, "Fault detection of uncertain nonlinear process using interval-valued data-driven approach," *Chem. Eng. Sci.*, vol. 205, pp. 36–45, Sep. 2019.
- [26] N. Huang, Q. Chen, G. Cai, D. Xu, L. Zhang, and W. Zhao, "Fault diagnosis of bearing in wind turbine gearbox under actual operating conditions driven by limited data with noise labels," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.
- [27] S. Yang, P. Yang, H. Yu, J. Bai, W. Feng, Y. Su, and Y. Si, "A 2DCNN-RF model for offshore wind turbine high-speed bearing-fault diagnosis under noisy environment," *Energies*, vol. 15, no. 9, p. 3340, May 2022.
- [28] F. E. F. Junior and G. G. Yen, "Particle swarm optimization of deep neural networks architectures for image classification," *Swarm Evol. Comput.*, vol. 49, pp. 62–74, Sep. 2019.
- [29] I. Ahmad, "Feature selection using particle swarm optimization in intrusion detection," *Int. J. Distrib. Sensor Netw.*, vol. 2015, no. 10, 2015, Art. no. 806954.
- [30] L. Hu, J. Zhang, Y. Xiang, and W. Wang, "Neural networks-based aerodynamic data modeling: A comprehensive review," *IEEE Access*, vol. 8, pp. 90805–90823, 2020.
- [31] Y. Liu, S. Liu, Y. Wang, F. Lombardi, and J. Han, "A survey of stochastic computing neural networks for machine learning applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 2809–2824, Jul. 2021.
- [32] C. Lv, Y. Xing, J. Zhang, X. Na, Y. Li, T. Liu, D. Cao, and F.-Y. Wang, "Levenberg–Marquardt backpropagation training of multilayer neural networks for state estimation of a safety-critical cyber-physical system," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3436–3446, Aug. 2018.
- [33] 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.
- [34] B. Hahn, M. Durstewitz, and K. Rohrig, "Reliability of wind turbines," in *Wind Energy*. Berlin, Germany: Springer, 2007, pp. 329–332.
- [35] J. Due, S. Munk-Nielsen, and R. Nielsen, "Lifetime investigation of high power IGBT modules," in *Proc. 14th Eur. Conf. Power Electron. Appl.*, 2011, pp. 1–8.
- [36] K. Dhibi, M. Mansouri, K. Bouzara, 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.





**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 research interest includes implementation of data-driven techniques for fault detection and diagnosis of industrial processes.



**MAJDI MANSOURI** (Senior Member, IEEE) received the degree in electrical engineering from the Higher School of Communications of Tunis (SUPCOM), Tunis, Tunisia, in 2006, the M.Sc. degree in electrical engineering from the National School of Electronics, Computer Science, Telecommunications, Mathematics and Mechanics of Bordeaux (ENSEIRB), Bordeaux, France, in 2008, the Ph.D. degree in electrical engineering from the University of Technology of

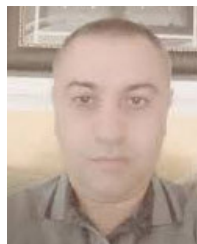
Troyes (UTT), Troyes, France, in 2011, and the H.D.R. (accreditation to supervise research) degree in electrical engineering from the University of Orleans, France, in 2019. He joined the Electrical Engineering Program, Texas A&M University at Qatar, in 2011, where he is currently an Associate Research 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.



**MOHAMED TRABELSI** (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the National Institute of Applied Sciences and Technology (INSAT), Tunisia, in 2006, and the M.Sc. degree in automated systems and the Ph.D. degree in energy systems from the National Institute of Applied Sciences of Lyon (INSA), Lyon, France, in 2006 and 2009, respectively. From October 2009 to August 2018, he has been holding different research positions at Qatar University and Texas A&M University at Qatar. Since September 2018, he has been an Associate Professor with the Kuwait College of Science and Technology. He has published more than 120 journals and conference papers, and he is the author of two books and two book chapters. His research interests include systems control with applications arising in the contexts of power electronics, energy conversion, renewable energies integration, and smart grids.



**KAMALELDIN ABODAYEH** received the M.Sc. degree in functional analysis from University College Dublin, the Ph.D. degree from University College Cork, Ireland, in 1997, and the Ph.D. degree from the Department of Process Engineering, University College Cork. Since 2001, he has been with Prince Sultan University, Saudi Arabia. He has published more than 60 articles in various areas of pure and applied mathematics. His research interests include functional analysis, theoretical physics, discrete potential theory, fixed point theory, and quality monitoring and statistical hypothesis testing.

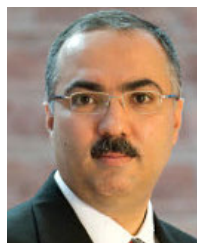


**KAIS BOUZRARAR** is currently a Professor of electrical engineering with the Laboratory of Automatic Signal and Image Processing, National Engineering School of Monastir, Monastir, Tunisia. He has more than 15 years of combined academic and industrial experience. He has published more than 80 refereed journals and conference publications and book chapters. His research interests include systems engineering and control, with emphasis on process modeling, monitoring, and estimation.



**HAZEM NOUNOU** (Senior Member, IEEE) is currently a Professor of electrical and computer engineering with Texas A&M University at Qatar. He has more than 19 years of academic and industrial experience. He has significant experience in research on control systems, database 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 for five

QNRF projects, some of which were in collaboration with other PIs in this proposal. He has published more than 200 refereed journals and conference papers and book chapters. He has served as an associate editor and on the technical committees for several international journals and conferences.



**MOHAMED NOUNOU** (Senior Member, IEEE) is currently a Professor of chemical engineering with Texas A&M University at Qatar (TAMU). He has more than 19 years of combined academic and industrial experience. He has successfully served as the lead PI and a PI for several QNRF projects (six NPRP projects and three UREP projects). He has published more than 200 refereed journals and conference publications and book chapters. His research interests include

systems engineering and control, with emphasis on process modeling, monitoring, and estimation. He is a Senior Member of the American Institute of Chemical Engineers (AIChE).

...