IUST
Iran University of Science and Technology

Iranian Journal of Electrical and Electronic Engineering

Journal Homepage: ijeee.iust.ac.ir

Research Paper

Fault Diagnosis of Wind Turbine Double-Fed Induction Generator Based on Multi-Level Fusion and Measurement of Back-to-Back Converter Current Signal

M. Kamarzarrin*, M. H. Refan*(C.A.), P. Amiri*, and A. Dameshghi*

Abstract: One of the major faults in Doubly-Fed Induction Generator (DFIG) is the Inter-Turn Short Circuit (ITSC) fault. This fault leads to an asymmetry between phases and causes problems to the normal state between current lines. Faults diagnosis from non-stationary signals for the Wind Turbine (WT) is difficult. Therefore, the strategy of fault diagnosis must be robust against instability. In this paper, a new intelligent strategy based on multi-level fusion is proposed for diagnosis of DFIG inter-turn stator winding fault. Firstly, to overcome the non-stationary nature of the vibration signals of the WT, empirical mode decomposition (EMD) method is performed in time-frequency domains to extract best fault features from information power sensor and information current sensor. Moreover, a feature evaluation technique is used for the input of the classifier to choose the best subset features. Secondly, Least Squares Wavelet Support Vector Machines (LS-WSVM) classifier is trained to classify fault types based on feature level fusion (FLF) from different sensors. The main parameters of SVM and the kernel function are optimized by Genetic Algorithm (GA). Finally, Dempster-Shafer evidential reasoning (DSER) is used for fusing the GA-LS-WSVM results based on decision level fusion (DLF) of individual classifiers. In order to evaluate the proposed strategy, a DFIG WT test rig is developed. The experimental results show the efficiency of the proposed structure compared to other ITSC fault diagnosis methods. The results show that the classification accuracy of DSER-GA-LS-WSVM is 98.27%.

Keywords: DFIG, DSER, EMD, Fusion, GA-LS-WSVM, ITSC.

1 Introduction

WIND turbine (WT) generators are often in two types Permanent Magnet Synchronous Generators (PMSG) and Doubly-Fed Induction Generator (DFIG) [1]. Induction generators operate in constant and variable speed modes. In the variable speed mode, the generator is connected to the grid from the stator side directly and from the rotor side indirectly (via a converter) [2, 3]. The under-study WT in this paper is a variable speed DFIG. WTs are largely installed in marine areas which are not easy to access them. Therefore, a continuous Condition Monitoring (CM) is necessary for reducing repair and maintenance costs [4, 5]. Generally, in induction machines including the DFIG, the Inter-Turn Short Circuit (ITSC) fault, which damages the insulation of windings, is common [6-8]. The ITSC detection prevents the expansion of the failure. Faults that are related to the stator winding insulation have weak signatures in signals and their detection is more difficult than another fault [8-11]. Moreover, the variable speed of the WT and its unstable dynamics cause disturbances and instability in electrical signals. Few researches related to ITSC fault in DFIG-based WTs have been conducted. However, many researches have been conducted about this fault in induction motors [10-14].

Iranian Journal of Electrical and Electronic Engineering, 2022.
Paper first received 04 January 2021, revised 04 October 2021, and accepted 16 October 2021.
* The authors are with the Faculty of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran.
E-mails: kamarzarrin.mehrnoosh@sru.ac.ir, refan@sru.ac.ir, pamiri@sru.ac.ir, and a.dameshghi@sru.ac.ir.
Corresponding Author: M. H. Refan.
https://doi.org/10.22068/IJEEE.18.2.2074

Forty percent of the electrical faults in WTs are related to the generator, and 40% of these faults are due to faults in the stator [6]. Investigating the faults related to stator shows that the fault of windings is the most important fault in stators. Damage to the insulator occurs due to ITSC, and partial discharge takes place between stator turns. The inter-turn fault emerges largely in the form of some kind of faults, such as coil-to-coil, phase-to-phase, open-circuit phase, and phase-to-ground [15]. This fault has little effect on the performance of the machine, and therefore, it can hardly be observed through performance parameters of the WT; however, its development is very costly and causes the destruction of the generator. In Table 1, an appropriate description of various conditions of fault in the stator windings is presented [9-15]. In the previous researches, the condition monitoring (CM) of WT based on vibration analysis technique has been proposed. The vibration analysis technique for ITSC fault is presented in [16, 17]. The vibration analysis technique requires adding sensors and related data collecting equipment. WT control system signals such as electrical signal cause a reduction in costs and eliminates the need for adding new sensors. In [18], two kinds of methods based on current are used. The first method is motor current analysis and Park's vector, the second method is a combination of wavelet transform and Park's vector methods. For a WT that does not operate in its stationary mode, using the second method is practical and reliable. In [19], as a result of investigation and analysis, it shown that the motor current analysis based on Park's vector cannot be used for a WT with a non-stationary signal. Therefore, the adaptive Park's vector method is used for detecting the ITSC in [6] and [20]. A method based on ITSC fault modeling for the DFIG is presented in [20]. The hybrid (ABC/dq) method is used for modeling the DFIG with ITSC fault [6]. In this reference, an index using the wavelet analysis is defined for ITSC fault; the proposed index was robust against changes in operating points, ITSC fault percentage, rotor/stator resistance, and the sampling frequency. In Ref [21], a method is presented based on measuring the rotor side transient current of the DFIG and voltage-pulse conversion for the asymmetry in rotor/stator phases. A method based on FFT analysis of reactive power signal is presented for detecting the ITSC fault in [22].The ITSC fault for a DFIG is detected by a simulation model based on finite element method [23]. In the healthy conditions of stator generator, the symmetry exists and the angle between phases is 120 degrees, the corresponding THD is low and the index of Park's vector trajectory is completely circular. Once the fault occurs, the THD and phase difference increase, and the circular shape of Park's vector trajectory undergo a distortion [24]. The basic model based on residual generation is proposed in [25] for ITSC fault detection. In this reference, the fault is modeled as a sensor using an observer. In [26], artificial neural network (ANN) is used for detecting and locating the ITSC fault in stator and rotor windings [26]. In [27], the inter-turn fault of rotor and stator winding is modeled using the magnetic equivalent circuit (MEC) model. The method of using the rotor current and coil voltage is presented in [28]. In [28], due to equation dependency and the proven mathematics in vibration and stator current signals, the use of rotor current is recommended. The multiple-coupled circuit model approach for stator fault in DFIG-based WTs is proposed in [29]. A dynamic DFIG state-space model is adopted in the presence of ITSC fault in [30]. In this reference, a Kalman observer is developed for fault detection based on residual generation. In [31], two types of stator winding faults in DFIG are investigated. These faults include ITSC and winding resistive asymmetrical faults. A complete MEC model is presented for modeling these two faults in stator winding of DFIG in [7]. Another paper based on MEC model is introduced in the [32]. In [33], a symmetrical components method is used for the detection of insulation faults in a three-phase induction motor.

In this paper, the main novelty is the presentation of a new Condition Based Maintenance (CBM) structure for detecting the stator fault of WT generator based on multi-sensor data fusion. Some fusion techniques for CBM are proposed in recent years. Only a few researches are conducted in the realm of WT fault diagnosis based multi-sensor fusion strategy. The fusion strategy through combining the qualitative and quantitative information leads to achieving an extensive and deep knowledge about faults. Fusion approaches are mainly intended for three purposes; i.e., data collaboration, data redundancy, and complementary data [34]. Three levels are considered for fusion: data or signal, feature and decision [35]. A method based on swarm intelligence and Bayesian inference is proposed for fusion of WT SCADA data in Ref [36]. The bearing fault diagnosis based on data fusion strategy for WTs is discussed in Ref. [37]. In Ref [38], WT gearbox CM and fault diagnosis are conducted based SCADA input signal and multi-sensor information fusion strategy. The

Table 1 Various conditions of fault in the stator windings.

	State	Action
1	inter-turn short circuit	Generator will continue to operate, but for how long?
2	Shorts between coils of the same phase	Generator can continue to operate but for how long?
3	Phase to phase short	Generator fails and protection equipment disconnects the supply
4	Phase to earth short	Generator fails and protection equipment disconnects the supply
5	Open circuit in one phase	Generator may continue to operate, depending on the load conditions

initial diagnosis is done by Relevance Vector Machine (RVM) and for the final decision-making is used of D-S evidential reasoning (DSER). For fault diagnosis in the planetary gearbox using the feature-level fusion and wavelet transform used of feature-extraction has been presented in another research [40]. An Artificial Neural Network (ANN) is used the fusion method in the feature level In Ref. [40]. For diagnosing the fault of the gearbox, the RVM method has been used for fusion in feature level of vibration sensors. Feature extraction is done by Ensemble Empirical Mode Decomposition (EEMD), and determining the main parameters of RVM is done by ant colony algorithm.

In this study, a hybrid (two-level) fusion strategy is presented for diagnosing the ITSC fault. Considering the better performance of Support Vector Machine (SVM), to ANN methods, this model is used for the fusion of feature vectors. In order to improve the efficiency of the model, three changes have been made to the conventional structure of the model. These changes include:

a) The higher computational load is the main drawback of SVM. This weakness is due to its optimization programming. The LS version of SVM simplifies the problem in such a way that the solution is characterized by a linear system. The Least Squares (LS) type of SVM model is used for this paper.
b) Due to the ability of wavelet functions to locally analyze and extract features, the conventional kernel function, which is Radial Basis Function (RBF), is replaced by a Wavelet (W) function.
c) Based on the practical application requirement of SVM, SVM, and the wavelet kernel parameters determination and optimization are conducted based on genetic algorithm.

Therefore in this paper, the first step electrical signal such as current and power is collected for fault diagnosis. In the second step, feature extraction from signals is conducted based on the time-frequency signal processing method. The measured electrical signal is usually a non-stationary signal. For feature extraction of a non-stationary signal, the time domain (RMS, crest, peak) or frequency domain (FFT) signal processing method is unsuitable. In this paper, the empirical mode decomposition (EMD) is used as a powerful signal processing method [41]. EMD decomposes a complicated signal into a set of intrinsic mode functions (IMFs). In this paper, time-domain and frequency-domain features are extracted from each IMF signal. In the third step, a distance evaluation technique [41, 42] is applied to reduce the redundant or irrelevant input features. In the next step, based on feature-level fusion (FLF), initial classification is performed, and finally, fault diagnosis analysis is finished by using fusion decision level fusion (DLF). The decision stage is based on Dempster-Shafer evidential reasoning (DSER) algorithm.

The paper is organized as follows; the methods of signal processing are introduced in Section 2. The optimized LS-SVM model with wavelet kernel for stator degradation classification is described in Section 3. Section 4 is fault diagnosis methodology. The case study and experimental setup are presented in Section 5. Section 6 is dedicated to discussing the results and comparison with other methods and the paper is finally concluded in Section 7.

2 Signal Processing

This section presents a brief discussion on the EMD signal processing method, feature Extraction, feature evaluation from the time domain, frequency domain, and time-frequency domain of current and power signals. The objective of this paper is to construct a new structure with high reliability for the diagnosis of WT generator ITSC fault, thus the three signals namely power output, converter stator-side three-phase current, and converter rotor-side three-phase current are chosen for signal processing. As stated, the ITSC fault is a result of damage to the insulation of the stator winding of the generator in areas between the coils [28]. Most of the condition monitoring and fault diagnosis methods in induction machines are based on current analysis. The capability of these methods of fault detection depends on the SNR ratio, the load on the machine, and most importantly, the performance of the machine in constant speed and stationary conditions [28-30]. In the stator current signal, Eq. 1 shows the frequency component of the fault in the corresponding spectrum [18].

$$f_{st} = f_1\left[n/p\,(1-s) \pm k \right] \tag{1}$$

where f_{st} is the frequency component, f_1 is the fundamental frequency, n = 1, 2, 3, … and k = 1, 3, 5. p is the pair pole of DFIG, fractional sliding is denoted by s. Based on this equation, the frequency component of the fault depends on sliding. In transient state and variability condition of WT s is varies [18-19]. Therefore, stator fault diagnosis based on the traditional signal processing method is difficult. In such conditions, it is better to use STFT or EMD for feature extraction, but STFT is dependent on the best selection of window, thus in this paper, EMS is used for processing of stator current, rotor current, and power signals. In the following, the presence signature of ITSC fault in rotor and power signals is discussed.

When exists an asymmetry in the stator-side of the DFIG, a negative harmonic component with frequency $-f_1$ produces a magnetic field in the stator. This magnetic field produces a harmonic component with frequency $(2–s)f_1$. If in the frequency domain, the ITSC fault is investigated in the rotor current signal, the index of this faulty component is the frequency $(2K±s)f_1$ [21, 28]. The stator winding has a voltage with constant amplitude. This voltage is directly connected to the grid and has a frequency equal to f_1. Once the ITSC fault in the stator occurs, regular positive and negative

sequences emerge, which their corresponding frequency are f_1 and $-f_1$, respectively. Therefore, the voltage of the grid or the voltage of the generator stator causes a new frequency component $2f_1$ in active and reactive power [22]. This frequency has amplitude two times larger than the power control loop and is easily observed in the signal. With regard to the fact that variations of the instantaneous reactive power are less than the instantaneous active power, reactive power signal is better for ITSC fault detection.

2.1 EMD

EMD method is used for processing non-stationary signal and operates the time-frequency domain. This technique performs the decomposition of the signal into a number of meaningful sub-signal components that representing a piece of information [41]. The instantaneous frequency and amplitude of each component can be estimated after decomposing the signal. The basis of the EMD method is the decomposition of the signal into a few oscillatory functions. These functions are called Intrinsic Mode Functions (IMF). The IMF function has to satisfy two conditions. The first condition is that in all data sets, the number of extrema and the number of passing through zero must be equal or has at most one unit difference. The second condition is that in every point, the average of the envelope defined by local maxima and the envelope defined by local minima must be zero. This implies the symmetry of IMF around zero. Consequently, for calculating the IMF the following procedure is performed [41, 42].

Step1: Determination of all the local extrema of the signal, $X(t)$ (including local maxima and minima of the signal).

Step 2: Interpolating the relative maxima and minima using the cubic-spline method for obtaining upper and lower envelopes of the signal.

Step 3: Calculating the average envelope from the upper and lower envelopes of the signal.

$$m(t)=\frac{X_L(t)+X_u(t)}{2} \tag{2}$$

Step 4: Subtracting the average envelope signal from the input signal.

$$d(t)=X(t)-m(t) \tag{3}$$

Step 5: Two conditions related to the IM function are checked; if they are satisfied $IMF_i = d(t)$ and the screening procedure finishes, otherwise the procedure returns to the first step.

Step 6: When the calculation of IMF in the fifth step finishes, the obtained IMF must be separated from the signal, and then, steps 1 up to 5 are performed on the remaining signals.

Step 7: If the remaining signal does not satisfy the two initial conditions, the procedure stops.

According to Steps 1-7, a signal with IMF is described as follows:

$$x(t)=\sum_{i=1}^{n} IMF_i(t)+r_n(t) \tag{4}$$

In Fig. 1, the flowchart of EMD is shown. In this paper, the first four IMFs have been used to get time and frequency domains features.

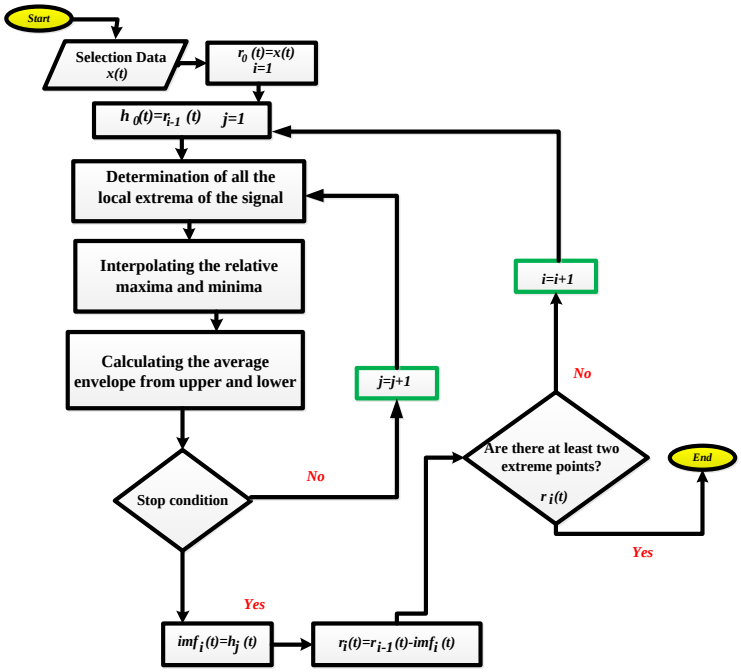


Fig. 1 EMD decomposition process.

2.2 Feature Extraction

When a fault occurs in WT, the behavior of electrical signals (i.e. current, voltage, and power signals) undergoes changes in the time domain. The amplitude and frequency spectrum are different from normal behavior for the time domain and frequency domain signals, respectively. In this paper, each of these signals is processed to extract ten time-domain features and ten frequency-domain features, as shown in Table 2. In this table, each feature based on energy, convergence, and position change of the main frequencies are categorized. Based on the EMD signal processing method, 20 indexes are defined for each IMF; therefore, each feature vector for each sensor contains 80 samples.

2.3 Best Subset of Features

In order to reduce the dimensions of the input matrix for a classifier algorithm, among the calculation features, a number of characteristics with appropriate criteria are chosen. The steps for choosing the best feature are as follows based on distance evaluation technique [43, 44]:

Step 1: The mean distance of data that has the same situations is calculated (Eq. (5)).

$$L_{a,b} = \frac{1}{n(n-1)} \sum_{m,g=1}^{n} \left| p_{a,b}(m) - p_{a,b}(g) \right| \tag{5}$$

In (5), $g \neq m$ and its value changes from 1 up to n, where n denotes the number of data from a specific feature with the same situations (the same fault percentage), $p_{a,b}$ denotes the value of the feature, $L_{a,b}$ is the mean distance of the same situations, a and b are the number of parameters and situations, respectively.

$$L_{Fa} = \frac{1}{M} \sum_{b=1}^{M} L_{a,b} \tag{6}$$

Table 2 Time-domain and frequency-domain features.

Time domain			Frequency domain		
Title	Equation	Description	Title	Equation	Description
$index_{t_1}$	$\frac{\sum_{n=1}^{N} x(n)}{N}$	Amplitude and energy	$index_{f_1}$	$\frac{\sum_{k=1}^{K} f(k)}{K}$	Energy
$index_{t_2}$	$\sqrt{\frac{\sum_{n=1}^{N} (x(n))^2}{N}}$	Amplitude and energy	$index_{f_2}$	$\frac{\sum_{k=1}^{K} (f(k) - index_{f_1})^2}{K-1}$	Convergence
$index_{t_3}$	$\left(\frac{\sum_{n=1}^{N} \lvert x(n) \rvert}{N} \right)^2$	Amplitude and energy	$index_{f_3}$	$\frac{\sum_{k=1}^{K} (sf(k) - index_{f_1})^3}{K(\sqrt{index_{f_2}})^3}$	Convergence
$index_{t_4}$	$\frac{\sum_{n=1}^{N} \lvert x(n) \rvert}{N}$	Amplitude and energy	$index_{f_4}$	$\frac{\sum_{k=1}^{K} f_k f(k)}{\sum_{k=1}^{K} f(k)}$	Position change of the main frequencies
$index_{t_5}$	$\frac{\sum_{n=1}^{N} (x(n))^3}{N}$	Amplitude and energy	$index_{f_5}$	$\sqrt{\frac{\sum_{k=1}^{K} (f_k - index_{f_4})^2 f(k)}{K}}$	Convergence
$index_{t_6}$	$\max(x(n)) - \min(x(n))$	Distribution	$index_{f_6}$	$\sqrt{\frac{\sum_{k=1}^{K} f_k^4 f(k)}{\sum_{k=1}^{K} f_k^2 f(k)}}$	Position change of the main frequencies
$index_{t_7}$	$\frac{\sum_{n=1}^{N} (x(n) - index_{t_1})}{N-1}$	Distribution	$index_{f_7}$	$\frac{\sum_{k=1}^{K} f_k^2 f(k)}{\sqrt{\sum_{k=1}^{K} f(k) \sum_{k=1}^{K} f_k^4 f(k)}}$	Position change of the main frequencies
$index_{t_8}$	$\frac{index_{t_2}}{index_{t_4}}$	Distribution	$index_{f_8}$	$\frac{\sum_{k=1}^{K} (f_k - index_{f_4})^3 f(k)}{k(index_{f_5})^3}$	Convergence
$index_{t_9}$	$\frac{\max(x(n))}{index_{t_2}}$	Distribution	$index_{f_9}$	$\frac{\sum_{k=1}^{K} (f_k - index_{f_4})^4 f(k)}{K(index_{f_5})^4}$	Convergence
$index_{t_{10}}$	$\frac{\max(x(n))}{index_{t_4}}$	Distribution	$index_{f_{10}}$	$\frac{\sum_{k=1}^{K} (f_k - index_{f_4})^{1/2} f(k)}{K(index_{f_5})^{1/2}}$	Convergence
$x(n)$: Signal series n: 1,2,…,N; N: Number of data.			$f(k)$: Spectrum k: 1,2,y,…,K; K: Number of spectrum lines; f_k: Frequency value of the k-th spectrum line.		

In (6), M represents different situations.

Step 2: The calculation of the mean distance between the data with different situations is performed.

$$L'_{Fa} = \frac{1}{M(M-1)} \sum_{m,g=1}^{n} \left| p_{Fa,m} - p_{Fa,g} \right|; \quad (m, g = 1,2\ldots,n; m \neq g) \tag{7}$$

where L'_{Fa} represents the mean distance between the data of different situations. Now $p_{Fa,b}$ as the mean data with identical situations is calculated as (8):

$$p_{Fa,b} = \frac{1}{n} \sum_{g=1}^{n} p_{a,b}(g) \tag{8}$$

Step 3: Selecting a type of feature in which this coefficient is larger.

$$\beta = \frac{L'_{Fa}}{L_{Fa}} \tag{9}$$

where β is a threshold, so the best features are selected based on the bigger distance from β coefficient.

3 FLL based on GA-LS-WSVM Classifier

ANN methods are based on the experimental risk minimization (ERM) and have some restrictions such as trapping in local minima, low convergence rate and low generalization capability [48]. SVM has structural risk addition to ERM and does not need much input data [49]. The structure of conventional SVM is similar to Fig. 2 [50]. The structure of SVM often encounters memory and time issues due to nonlinear solution and optimization of the main equations [48-50]. In order to solve these issues, LS-SVM, which uses linear equations instead of quadratic programming, is utilized. Furthermore, the SVM has parameters that their optimization hugely influenced in SVM performance [51]. There are three common solutions for this purpose: using grid search, GA, and Particle Swarm Optimization (PSO) [51, 52]. In this paper, instead of using the conventional kernels, such as linear kernel, Polynomial, RBF, and MLP, the wavelet transform kernel is used.

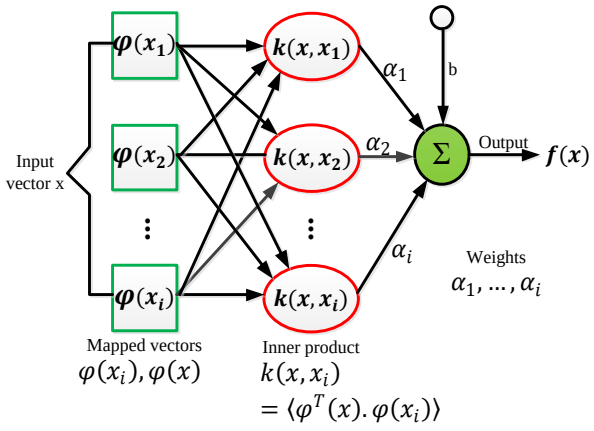


Fig. 2 Conventional structure of SVM.

3.1 LS-SVM

By supposing a training set in the form $\{(x_i, y_i) \mid x_i \in R_n, y_i \in \{-1, +1\}\}^N_{i=1}$, where x_i is the input and y_i is the binary class label, the discriminant function is extracted according to (10) [53].

$$y = sign\left[w^T \varnothing(x) + b \right] \tag{10}$$

In (10), $\emptyset(0)$ is a nonlinear function that maps the input to the high-dimensional feature space. This function is not explicit and may have infinite dimensions. Moreover, in (10), the vector w and the bias term b are to be determined. For determining these two terms of (10), the following optimization problem must be solved:

$$\min J(w,e) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{i=1}^{N} e_i^2 \tag{11}$$

where γ is a parameter for regulation. Considering the following equation:

$$y_i \left[w^T \varnothing(x_i) + b \right] = 1 - e_i, \qquad i = 1.2,\ldots,N \tag{12}$$

The Lagrangian corresponding to (11) is defined by (13).

$$L(w,b,e_i,\alpha) = J(w,b,e) - \sum_{i=1}^{N} \alpha_i \left\{ y_i [w^T \varnothing(x_i) + b] - 1 + e_i \right\} \tag{13}$$

where αi is the Lagrange multiplier. The KKT conditions are defined as follows:

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{N} \alpha_i y_i \varnothing(x_i) \\ \dfrac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{N} \alpha_i y_i = 0 \\ \dfrac{\partial L}{\partial e_i} = 0 \rightarrow \gamma \alpha_i \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i \left[w^T \varnothing(x_i) + b \right] - 1 + e_i = 0 \end{cases} \tag{14}$$

In LSSVM, in contrast to the conventional SVM, solving (11) is not performed by QP solution. A Linear solution of (11) is performed according to the following equation:

$$\begin{bmatrix} 0 & y^T \\ y & ZZ^T + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{1} \end{bmatrix} \tag{15}$$

where $Z = [\emptyset(x_1)y_1, \ldots, \emptyset(x_N)\, y_N]^T$, $y = [y_1, y_2, \ldots, y_N]^T$, $\vec{1} = [1, 1, \ldots, 1]$, $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_N]$. We have

$$\mu = ZZ^T \tag{16}$$

$$\mu_{kl} = y_k y_l \varnothing(x_k)^T \varnothing(x_l) = K(x_k, x_l) \quad (k,l = 1,2,\ldots,N) \tag{17}$$

Finally, after solving (14), the LS-SVM for classification will be obtained as follows:

$$y = \operatorname{sign}\left[\sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + b\right] \tag{18}$$

3.2 Wavelet Function as Kernel LS-SVM

The type of kernel function is important. As a general rule, if a function satisfies the Mercer condition, it can be used as a kernel function; it has been proven that the wavelet transform satisfies this condition [54]. If the wavelet function $\psi(x)$ satisfies $\psi(x) \in L^2(R) \cap L^1(R)$ and $0 = \psi(x)$, then $\overline{\psi(x)}$ is the Fourier transform of $\psi(x)$. This function is a time-frequency analysis. Also, this function can be recovered from the main signal. The method of inserting this function as the LS-SVM kernel function is presented in [55]. Assuming that $\int_0^{\infty} \frac{|\psi'(x)|}{\omega} d\omega < \infty$ and if $\psi(x)$ is on the path of realization of Fourier transform, and then $\psi(x)$ is called the mother wavelet. Therefore, the wavelet functions group is defined as the following equation:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \tag{19}$$

where a, $b \in$ R, and a is larger than zero. A multidimensional wavelet function is made based on a one-dimensional wavelet function.

$$\psi(x) = \prod_{i=1}^{N} \psi(x_i) \tag{20}$$

The wavelet function that is inserted into SVM is according to (21).

$$K(x, x') = \prod_{i=1}^{N} \psi\left(\frac{x_i - x_i'}{a}\right) = \prod_{i=1}^{N} \left[1 - \frac{(x_i - x')^2}{2}\right] \exp\left(-\frac{(x_i - x')^2}{2a^2}\right) \tag{21}$$

After finding the kernel function, the final LMS-WSVM function is obtained as (21):

$$y = \operatorname{sign}\left[\sum_{i=1}^{N} \alpha_i y_i \prod_{i=1}^{N} \psi\left(\frac{x_i - x_i'}{a}\right) + b\right] \tag{22}$$

3.3 Design of FLL based on LS-WSVM

Since there are various types of faults in this paper, the multiclass SVM is used for the FLL procedure. Among different kinds of multiclass SVMs, the one-against-all method is used in this paper. For k classes, a number of k binary SVMs must be utilized [56]. For estimating the posterior probability corresponding to the SVM binary, a sigmoid function is used, for details about this function). Finally, the posterior probability corresponding to the multiclass SVM binary is obtained from the following equation:

$$P_i(x) = P\left(y = i \mid d_i(x)\right) = \begin{cases} \dfrac{1}{1 + \exp\left((\ln u)\, d_i(x)\right)} & |d_i(x)| < 1 \\ 1 & |d_i(x)| \geq 1 \end{cases} \tag{23}$$

while $P_i(x)$ is the posterior probability of x belonging to class I and $d_i(x)$ is the normalized distance of x with respect to the line separating the SVM binary i. The separating hyperplane is considered as normalized and it is between zero and one.

$$p_i(x) = \frac{P_i(x)}{\sum_{i=1}^{k} P_i(x)} \tag{24}$$

Therefore, the following equation holds for a multiclass SVM.

$$D(x) = \arg\max\left\{p_1(x), p_2(x), \ldots, p_k(x)\right\} \tag{25}$$

In this paper, a number of L SVM multi-classifiers exist that are used for k classes. The decision making a profile for the analyses corresponding to decision level fusion is given as follows:

$$DP(x) = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \cdots & \cdots & \ddots & \vdots \\ p_{L1} & p_{L2} & \cdots & p_{LK} \end{bmatrix} \tag{26}$$

where p_{ij} is the posterior probability of x belonging to the i-th, $i = 0, \ldots, k$ class in the j-th, $j = 0, \ldots, L$ classifier.

3.4 Optimization of Parameters

Since wavelet is chosen as the kernel function, the proper performance of SVM depends on two parameters γ and a. Choosing a value for γ less than the proper value causes an unbalance between model complexity minimization and ERM. If the value of a is selected smaller than the proper value, then SVM will experience an accumulation of input data. If a is selected large than the proper value, then WSVM will

not solve complex problems with flexibility as much as needed. One of the methods that are proposed for determining these parameters is the GA [52]. The principal concepts of the genetic algorithm include three operators, i.e., selection, crossover, and mutation. The main steps of a genetic algorithm are shown in Fig. 3. The free parameters, γ and C are represented by a chromosome in Fig. 4 in the form of binary numbers. Here the size of the population is set to around 200, meaning the inclusion of 40 bits by each gene. Parents are selected using two chromosomes from the entire population based on the fitness function. For selecting chromosomes in reproduction, the roulette wheel selection is used. In this operation, the chromosomes are paired randomly. Some pairs of chromosomes that are selected between two breakpoints are shifted. Before the crossover is done, the values of the two parameters for the first parents are 1.125 and 3.125; for the second parents these values are 0.375 and 8.75. After crossover the values of the two parameters for the first offspring are 1.375 and 3.75; for the second offspring these values are 0.125 and 8.125. A mutation is performed randomly by converting a zero bit into a one bit or vice versa. The rates of crossover and mutation are determined in a probabilistic manner. Fig. 5 is the genetic process of determining the main parameters of WSVM.

4 Fault Diagnosis Methodology

The main idea of the proposed structure is to use data fusion of several sensors. Fusion strategy has different levels. In this paper, fusion in FLF and DLF are addressed. The structure of fault diagnosis methodology based on fusion strategy is shown in Fig. 6. The structure includes five stages. In each stage, some functions are used. In the first stage, data collection is performed. In the second stage, the signal processing method is implemented. The third stage includes feature extraction and selection in both time and frequency domains. Fusion in feature level is done in the fourth

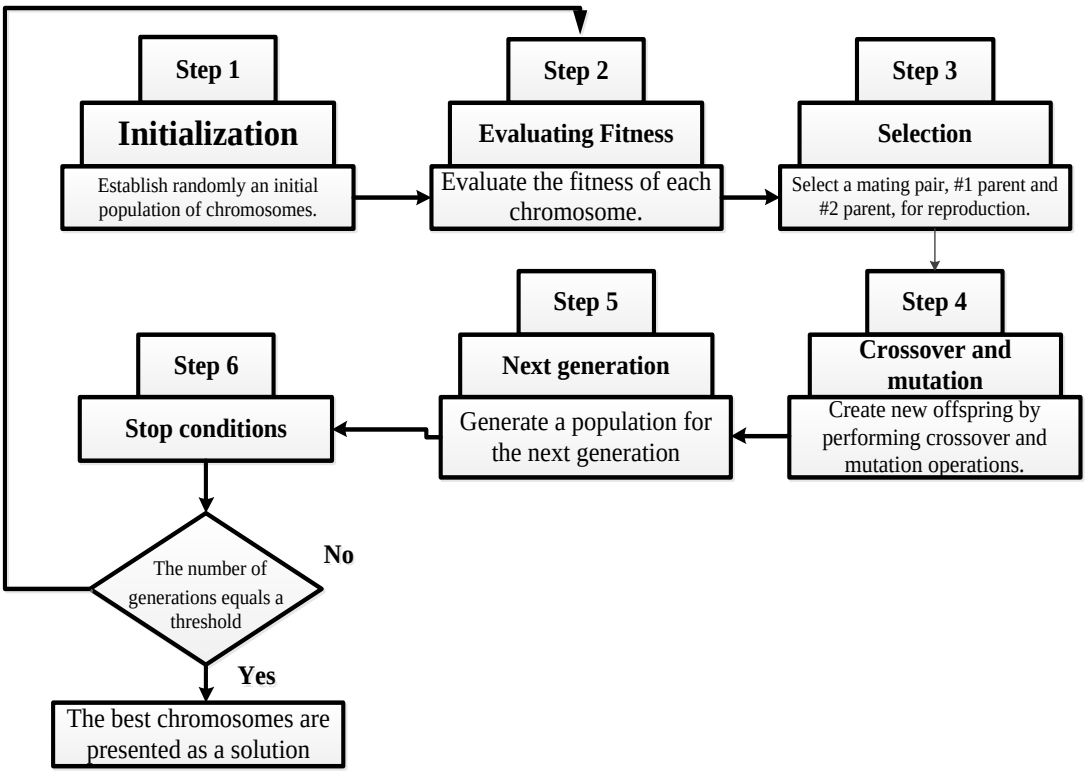


Fig. 3 Main steps of genetic algorithm.

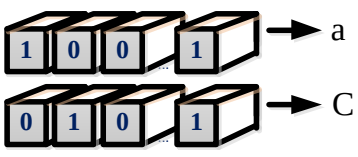


Fig. 4 Gene mapping of the main parameters of WSVM.

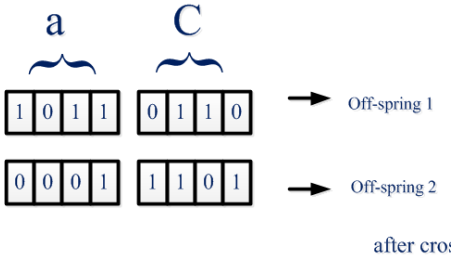


Fig. 5 Genetic process for determining the main parameters of WSVM.

stage, and finally, fusion in decision level and fault diagnosis is performed in the fifth stage. The structure diagram of the ITSC fault diagnosis based on FLF is shown in Fig. 7. Three profiles decision are used for final decisions about the type and intensity of the fault. Using a single-sensor data source was common for condition monitoring and fault diagnosis in the prior research and commercial solution. For complicated machinery such as a WT, a single sensor is not capable of covering all features of a fault. Using a multi-sensor scenario based on the fusion strategy is the solution to the issue. This solution leads to the best decision. In this paper, two fusion levels are used in a combinatory form. The feature of each data sensor is extracted individually, and then, these features gathered in a group are used as the input for the algorithm. In fault diagnosis applications, the data from homogeneous sensors are fused in the feature and decision levels, and the data from heterogeneous sensors are fused at the decision level.

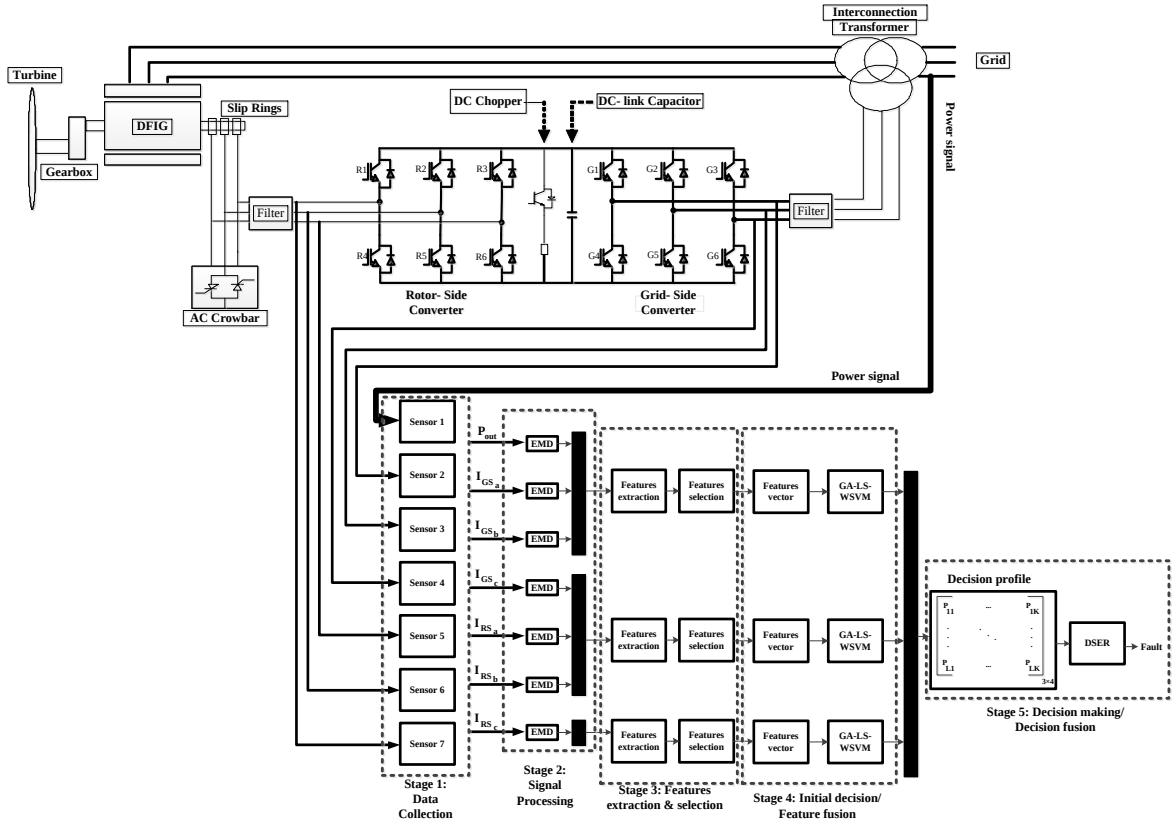


Fig. 6 Block diagram of the proposed approach.

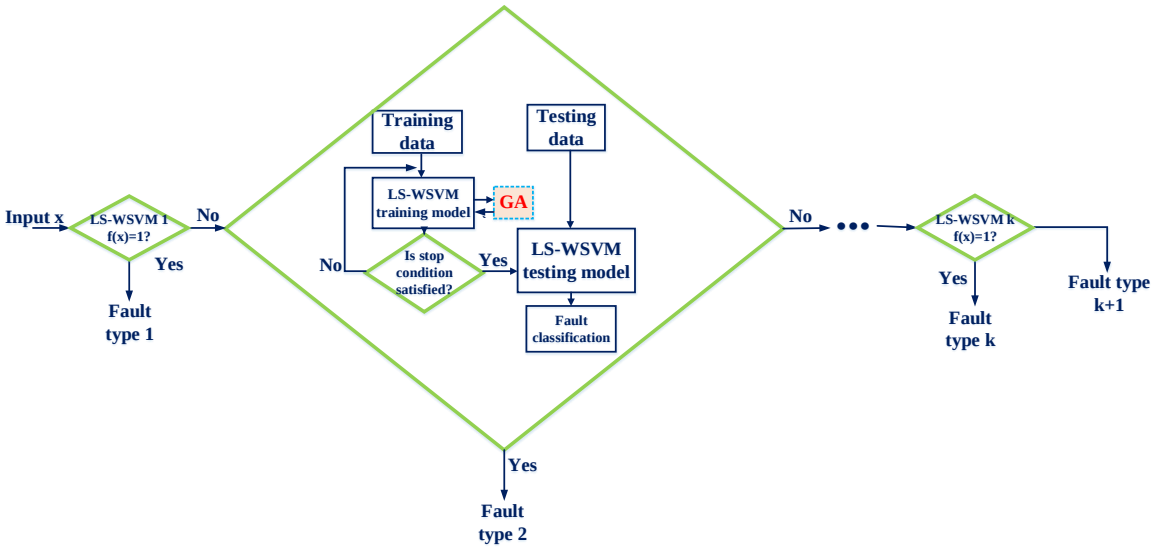


Fig. 7 Block diagram of the multi class GA-LS-WSVM.

4.1 Decision Level Fusion (DLF)-DSER

By assuming that a set according to (27) is a countable set of probabilities belonging to the realization of one possibility, this is a power set denoted by 2^{Θ} and the framework of investigations is about the occurrence of an event. The sample space is called a frame of discernment, and $H_j \in \Theta$ [57].

$$H = \{H_1, H_2, \ldots, H_k\} \tag{27}$$

The Dempster–Shafer theory consists of three important and fundamental functions: probability mass function, belief function and plausibility function [58]. The mass function is described as follows:

$$\begin{cases} m: P(X) \to [0,1] \\ m(\varnothing) = 0 \\ \sum_{A \in P} m(A) = 1 \end{cases} \tag{28}$$

In the above equation, $P(X)$ denotes the power of the evidence X and shows all probabilities belonging to the realization of one possibility. The symbol ∅ represents an empty function and A is a possible state from the power set. Therefore, the mass function indicates mapping the evidence to the occurrence of state A. This mapping is stated by a value between zero and one, provided the probability of the empty state is zero and the sum of mass functions of all subsets of the reference set is equal to unity. The value of believing in the occurrence of set A ($Bel(A)$) is the sum of mass functions of set B, which is the subset of A. In fact, the belief function is the pessimistic estimate of a possible state such as A. The plausibility or PLA is obtained by summing over mass functions of the set B (so that $B \cap A = \emptyset$).

$$Bel(A) = \sum_{B|B \subseteq A} m(B) \tag{29}$$

$$PL(A) = \sum_{B|B \cap A = \varnothing} m(B) \tag{30}$$

According to these equations, it is obvious that Bel(A) ≤ $PL(A)$. Since the belief and plausibility of A are calculated from its mass function, it can be written:

$$PL(A) = 1 - Bel(A) \tag{31}$$

The probability of some event A, denoted by $P(A)$, is a value between the plausibility and belief of that event. Therefore, the following relation holds:

$$Bel(A) \le P(A) \le PL(A) \xrightarrow{\text{if } PL(A)=Bel(A)} PL(A) = P(A) = Bel(A) \tag{32}$$

Combination rules are the foundation of Dempster–Shafer evidence theory. The calculation of belief and plausibility functions is performed based on the combination of mass functions. In this method, the combination of evidence theories is performed based on AND operator, and the sum of probability of an event such as A is calculated by multiplying the probabilities of all evidence by the probability of events that are equal to or including A. Combination of two and multiple mass functions are done by (33) and (34). In these equations, all the sets that do not intersect are rejected by the factor K in order to prevent them from influencing the results of estimates, and therefore, $NF = 1/(1–k)$ is called the normalizing factor.

$$m_{12}(A) = m_1(B) \oplus m_2(C) = \frac{\sum_{B \cap A = A} m_1(B) m_2(C)}{1 - \sum_{B \cap A = \varnothing} m_1(B) m_2(C)} \tag{33}$$

$$m_1(A) \oplus m_2(A) \ldots \oplus m_i(A) = (1-K)^{-1} \times \sum_{\cap A_i = A} m_i(A_i) \tag{34}$$

The output value of each classifier is unique; the degree coefficient, $d_i (j = 1, 2, \ldots, L)$, is applied to BPA before the combination of evidence. It is assumed in this paper that $M_i(H)$ is the main BPA of the classifier j. Therefore, the revised quantities are defined by the following equations:

$$m_i(H) = d_i M_i(H) \tag{35}$$

$$m_i(\theta) = 1 - d_i \tag{36}$$

Finally, if r-evidence exists, the D-S decision-making theory for fusing the classifiers is defined as

$$\begin{cases} m_r(H_{k0}) - m_r(H_{k1}) > \tau_0 \\ m_r(\theta) < \tau_1 \\ m_r(H_{k0}) > m_r(\theta) \end{cases} \tag{37}$$

where $m_r(H_{k1}) = \max\{m_r(H_j), H_j \neq H_{k0}\}$, $m_r(H_{k0}) = \max\{m_r(H_j), H_j \subseteq H_j = 1, 2, \ldots, k\}$, τ_0 and τ_1 are the threshold values, which are obtained experimentally or under specific rules. In this paper, threshold values are regulated based on experimental data.

4.2 Fault Diagnosis Steps based on FLF and DLF

On the basis of the FLF (Fig. 6) and DLF (Fig. 7), the processes of ITSC fault diagnosis are described below:

Step 1: The measurement signal is carried out through test rig. The electrical signals include stator currents, rotor currents, and power are acquired with acquisition systems.

Step 2: EMD signal processing method is applied on the signals. The collected electrical signals are decomposed into n IMFs.

Step 3: The four first IMFs are selected and the features extraction is conducted based on time-domain and frequency-domain.

Step 4: Best subset of features is selected based on (5)-(9).

Step 5: The feature vectors profiles are made and these vectors are divided into training and testing sets.
Step 6: Run the LS-SVM.
Step 7: The kernel function is selected.
Step 8: The GA algorithm is executed.
Step 9: The multi-class GA-LS-WSVM classifier is produced based on Fig.6. Initial diagnosis of each is obtained using (24)-(26).
Step 10: The DLF stage is conducted. DSER is applied for decision-level fusion based on the decision profile. In this stage, the results of the final diagnosis are obtained.

5 Experiment Setup

5.1 The Case Study Fault

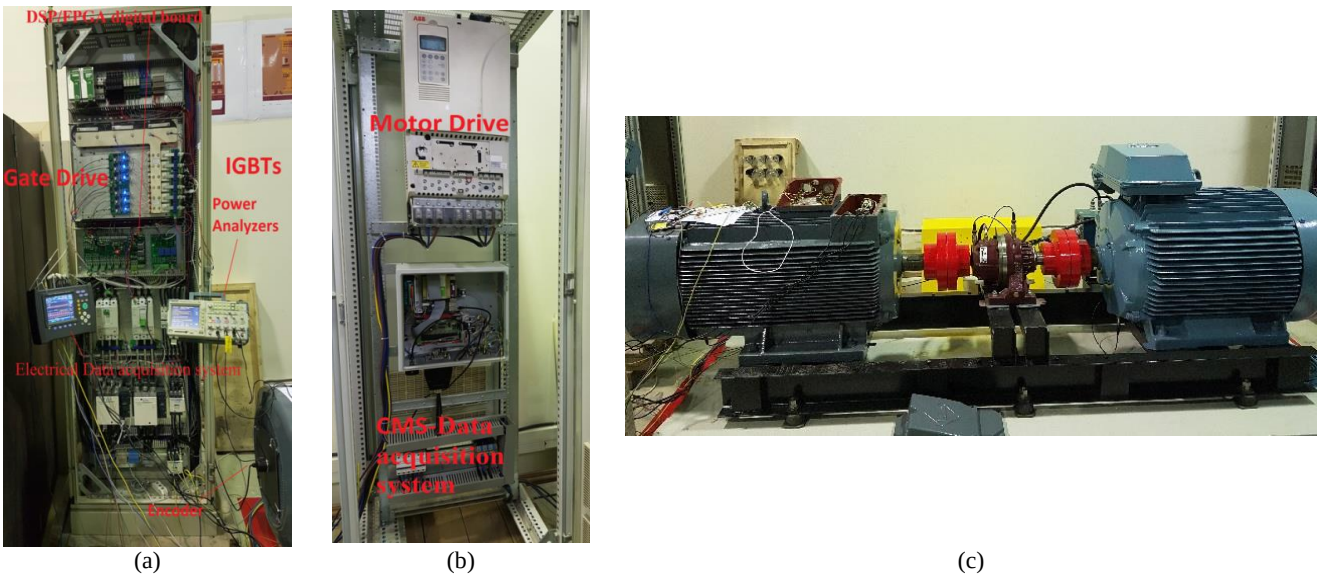
With regard to the fact that the ITSC fault occurs due to damages to the insulation between the turns of a winding, it is modeled as a short circuit in the winding that affects the stability of current. This modeling is according to Fig. 8(a). For this purpose, it is assumed that in the modeling, the impedance of the short-circuited stator winding is reduced. The amount of this reduction is related to the fault severity. Hence, in MATLAB/Simulink, a variable resistance, similar to Fig. 8.b is placed in parallel to one of the phases. The number of short-circuited turns is determined according to the changes in this resistance [19-22].

5.2 WT Test Rig

The verification of the proposed method is done on a designed WT test rig by the authors, that shown in Fig. 9; the test rig characteristics are shown in Table 3. This test rig is used to test converter fault and it has communication with the CM module. The ITSC is applied under different operating conditions of WT. This test rig contains a 90-kW DFIG coupled to a sun gearbox with a transmission 1 to 3.33. The gearbox is rotated by a 110-kW asynchronous motor. The 8-poles DFIG operates at 400 V and 50 Hz. The system comprises two voltage source inverters in back-to-back converter topology with a high power rate 30 kw. In this topology, there are gate drive circuit to trigger of 12 gates of IGBT and DFIG crowbar circuit to protection of short circuit current. An ARM/FPGA digital board with 100 kHz sampling frequency is used for the control system of the converter. The control of the converter is based on vector control strategy and is doing by FOC and GOC with the SVM method. For measuring mechanical speed, an encoder that is coupled to the shaft of the DFIG generator is used. Wind pattern emulated to ABB-ACS800-31 motor drive using the torque-speed curve is based on the LabView interface and using a PC. The test rig works under “healthy” and “faulty” conditions during the experiments. Relevant signals are collected from the generator and the

(a) (b)

Fig. 8 The structure of stator winding with ITSC fault in phase C.



(a) (b) (c)

Fig. 9 DFIG wind turbine test rig.

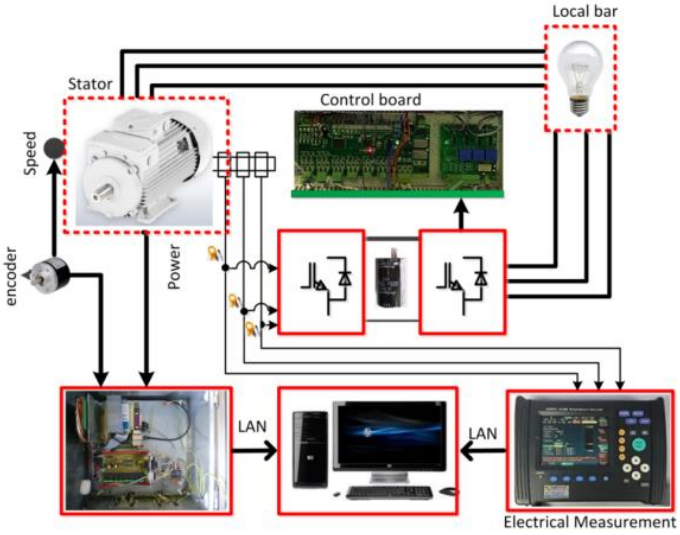


Fig. 10 Data acquisition system configuration.

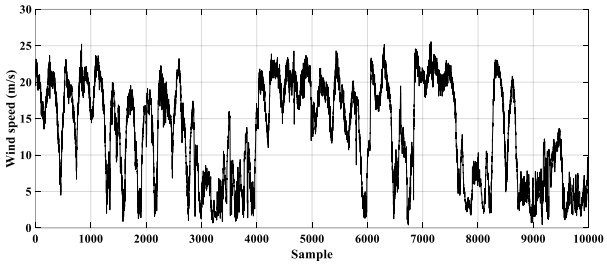


Fig. 11 Real wind speed data emulated on test rig.

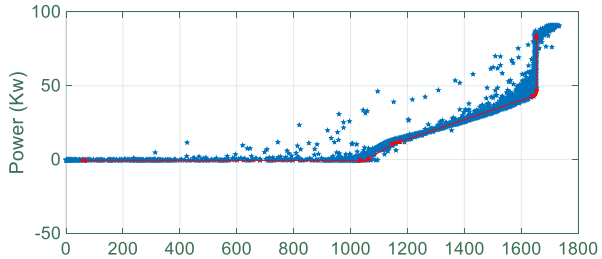


Fig. 12 Power output-generator speed curve for DFIG test rig.

converter. The different types of electrically asymmetric fault are applied on the stator, representing the effect of an ITSC fault in the WT generator. Two systems for data collection in this test rig are used. A data acquisition system was used to obtain rotor and stator currents. Another system is WT CMS. This system is used to measuring of operational data (power, generator speed, and wind speed) of WT. An Overview of the data acquisition system and communication is presented in Fig. 10.

The Hardware simulator input information including wind speed data was collected from Kahak Qazvin power plant. The input to the simulator is the real wind speed data collected by an anemometer from a MAPNA's wind farm located in Takestan-Iran. The wind speed data includes different operating ranges (cut-in, transition, power generation, and cut-out) from less than 1 m/s up to 28 m/s. The data leads to the operation of the DFIG in sub-synchronous, synchronous the super-synchronous ranges. The wind speed information (Fig. 9) is used as input to the drive of the propulsion motor. The torque reference is applied to the motor and rotates in the various operating areas of the WT. Fig.10 shows the power curve of the 90 DFIG WT test rig. The unstable dynamics of WT and WT operation at variable speed, which is due to instability and turbulence in the measurement signals, are well described in Figs. 9 and 10.

Table 3 Parameters of test-rig components.

Drivetrain	
Power	90
Speed	1488 RIV/MIN
Stator voltage	400 V
DC link voltage	700 V
Current	199
Torque	578 Nm
Number of pole pairs	4
Magnetizing inductance	120.4 mH
cos φ	0.88
Stator resistance	24.8 mΩ
Stator inductance	44 mH
Rotor resistance	16.6 mΩ
Rotor inductance	33 mH
Gearbox	Gear with 1 stage by 1:3.32
Prime mover motor	8 pole / 55 kW/ 400 V/ 742 rpm
Converter	
Filter resistance	5 mΩ
Filter inductance	800 μH
Sampling frequency	100 kHz
Switching frequency	2.5 kHz
Grid side converter	FOC
Rotor side converter	GOC

Table 4 Phase difference between three phases of stator in the presence of fault.

Fault %	AB	BC	CA
Normal	120.012	120.063	119.358
5 %	122.747	120.087	116.996
10%	125.536	120.997	113.467
25%	135.096	119.667	105.237
40 %	142.198	119.558	98.244

Table 5 Variation of THD in various fault percentages.

Fault %	A	B	C
Normal	3.39%	3.54%	3.20%
5 %	3.62%	3.57%	3.22%
10%	3.92%	3.65%	3.18%
25%	5.59%	3.52%	3.23%
40 %	7.98%	3.48%	3.39%

6 Experimental Result and Discussions

6.1 Experimental Result

In order to study the relations between the fault features and the fault degree and training the two-level fusion algorithm, various test data are needed. This experiment is performed for analyzing the behavior of ITSC fault. In this section, various tests are designed in stationary conditions. Every percent of asymmetry is proportional to the coil turns. In Table 4, for different percentages of fault in phase A, the phase difference

Table 6 Numerical analysis of trajectory of Park's vector.

Fault %	Eccentricity
Normal	0.010
5 %	0.251
10%	0.348
25%	0.518
40 %	0.620

between three phases of the stator are shown. As the fault percentage in phase A increases the phase difference between A and B phases increases, while the phase difference between two healthy phases has a very small change (close to 120 degrees). In Table 5, the variation of THD is presented. In this table, it is seen that the fault percentage increases with THD, but in the other two healthy phases, the variation of fault percentage is small and close to THD in healthy state. Transformation of three phases into two phases is performed by Park's transformation. In equilibrium state, the trajectory of Park's vector is circular, and when the ITSC is exerted, the trajectory becomes elliptical. Table 6 is a numerical analysis of the trajectory of Park's vector. For Investigating the accuracy and precision of the proposed structure, the ITSC fault scenario is randomly considered as Table 6. The data collecting operation is performed each 1 ms. With regard to real wind speed (Fig. 11), all operation ranges of WT from sub-synchronous to super-synchronous are taken into account. Therefore, the collected data includes both stationary and transient ranges. The features extracted in the fusion-based fault diagnosis process are very important because the real results of the proposed structure are obtained when a proper data collecting is done and the features corresponding to the fault are extracted. In this paper, 80 features have been defined for each signal. In order to reduce the dimensions of the input matrix for the fusion algorithm, the best feature are selected based on the β coefficient (β is selected 2.8). As result, the first, second and third fusion blocks in the feature level contain 100, 300, and 300 input samples, respectively.

Among total input features, 60% are used for training and 40% for testing. In this paper, three blocks of GA-LS-WSVM classifier are used for seven electrical sensors. For performance evaluation of GA-LS-WSVM, this method is compared with its other family members, i.e., the SVM with RBF kernel function, the SVM with a nonlinear solution and without LS and the SVM optimized with PSO. The precision of the classification is calculated based on the parameter CA according to (38). Considering the fact that there are five classes, the average precision is calculated from (39) and is called ACA.

$$CA_i = \frac{H_c^i}{H_t^i} \times 100 \tag{38}$$

$$ACA = \frac{H_c}{H_t} \times 100 \tag{39}$$

In the above equations, H_t^i is the total number of the samples from class i and H_c^i is the number of samples that are properly classified. Moreover, H_t is the total number of samples and H_c is the total number of samples that are properly classified. Investigation and statistical analysis are performed separately for different stages of testing and training data. In the training stage, the optimal parameters of kernel are chosen proportional to the kernel function. In the analysis section, there are three classifiers and five classes; in fact, the decision profile consists of a matrix with five rows and three columns. In this paper, there are three SVM multiclass classifiers, and DSER exists in the fusion-decider section. Therefore, the methods GA-LS-WSVM-DSER, PSO-SVM-DSER, and GA-LS-RBF-SVM-DSER are utilized for ITSC fault diagnosis and are compared with the case in which they are used individually without DSER. DSER operates based on H_1–H_5 states, which are the ITSC fault diagnosis classes. There are three evidences for combining and $r = 3$. Before fusion, BPA has to be calculated and for its calculation, the decision profile matrix, which has five rows and three columns, must be taken into account. For each evidences, a coefficient is defined. In this paper $d_j (j = 1, 2, 3)$, and therefore, these coefficients are 0.85, 0.75, and 0.70 for power signal, three phases of stator and three phases of rotor, respectively. In this research the experimental values of τ_0 and τ_1 are 0.001 and 0.04, respectively. Three analyses are discussed:

1. The performance of each individual classifier indicates the fusion only in the FLF without combining the ideas.
2. The performance of classification and fault diagnosis based on FLF and decision DLF in a hybrid form.
3. Fault diagnosis based on each signal separately and without performing fusion strategy.

Tables 7-9 show the results of analyses 1-3, respectively. Comparing Tables 7 and 8 reveals that ACA factor for individual classification in testing and training conditions are 90.09% and 95.88% for PSO-RBF-SVM, 94% and 97.016% for GA-LS-WSVM and 87.23% and 91.86% for RBF-SVM, respectively; while, in fusion mode ACA of three methods for the testing set and training set are 95.51% and 97.83% for DSER-PSO-SVM, 98.27 and 99.36 for DSER -GA-LS-WSVM and 90.34 and 93% for DSER-RBF-SVM. This implies that combining in decision level has increased the accuracy of fault diagnosis. Moreover, in both tables, the performance of GA-LS-WSVM method is better than other classifier algorithms in the same family, and this implies that the wavelet kernel function is better than RBF. If the same algorithm is used alone and the features extracted from the seven signals are used directly without fusion in the feature level as inputs for GA-LS-WSVM, then the accuracy will be according to Table 9. It can be seen that the accuracy of the fault diagnosis is noticeably less than the values in

Table 7 Results of analyses with only FLF.

Class	Fault	CA_i	PSO-RBF-SVM		GA-LS-WSVM		RBF-SVM	
			Test	Train	Test	Train	Test	Train
1	F_1	CA_1	90	96.23	95	98.2	85	93.44
2	F_2	CA_2	91	95	94	97.3	88.65	91
3	F_3	CA_3	90.23	94.35	93	96.23	90.16	93.48
4	F_4	CA_4	89.23	96.83	95	98.35	83.23	89.16
5	Normal	CA_5	90	97	93	95	88.99	92.26
	Total	ACA	90.09	95.88	94	97.016	87.23	91.86

Table 8 Results of analyses with FLF and DLF hybrid fusion.

Class	Fault	CA_i	DSER-PSO-SVM		DSER -GA-LS-WSVM		DSER-RBF-SVM	
			Test	Train	Test	Train	Test	Train
1	F1	CA_1	97.32	100	100	100	90.23	93
2	F2	CA_2	96.82	98	98.82	100	89.66	94
3	F3	CA_3	98.23	100	100	100	93.73	95
4	F4	CA_4	95.95	97.04	95.81	98.8	86.89	90
5	Normal	CA_5	89.23	94.11	96.74	98	91.23	93
	Total	ACA	95.51	97.83	98.27	99.36	90.34	93

Table 9 Results of analyses without fusion and based on seven signals with GA-LS-WSVM algorithm.

Class	Fault	CA_i	I_{GS1}		I_{GS2}		I_{GS3}		I_{RS1}		I_{RS2}		I_{RS3}		P	
			Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train
1	F_1	CA_1	56	63	57.63	61.14	60.00	64.21	55.21	62.23	56.83	59.83	62	62.11	71.23	75
2	F_2	CA_2	61.04	65	58.21	62	61.21	63.27	52.63	59.98	58.18	60	60	65.14	69.23	73
3	F_3	CA_3	62.82	66	59.16	64	60.21	65.12	59.18	66.12	57.12	60.89	61.14	63.04	73.41	76
4	F_4	CA_4	55.23	61	57.48	71	60.12	62.29	56.01	63.73	54.12	65	58.12	62	75.84	81
5	Normal	CA_5	58	62	58.21	69.18	60.09	69.12	57	68	52.22	61.45	61.12	67	76.32	80
	Total	ACA	58.61	63.40	58.13	65.46	60.32	64.80	56.00	64.01	55.69	61.43	60.47	63.85	73.20	77

Table 10 Results of analyses with only feature level fusion.

Class	Fault	CA_i	BPNN		GA-RNN	
			Test	Train	Test	Train
H_1	F_1	CA_1	69.86	75	73.66	83.16
H_2	F_2	CA_2	61.04	69.78	71.12	80.01
H_3	F_3	CA_3	73.12	78.46	72.13	80.78
H_4	F_4	CA_4	70.01	75.05	75.14	83.12
H_5	Normal	CA_5	69	73.23	78.12	86.18
H	Total	ACA	68.60	74.30	74.03	82.65

Table 11 Results of analyses with hybrid feature and decision level fusion.

Class	Fault	CA_i	DSER-BP-RNN		DSER -GA-RNN	
			Test	Train	Test	Train
H_1	F_1	CA_1	78.29	85.21	80.11	83.20
H_2	F_2	CA_2	73.11	87.33	74.11	75.66
H_3	F_3	CA_3	72	89.21	73.12	75
H_4	F_4	CA_4	76.10	76.12	77	78.26
H_5	Normal	CA_5	77	78.11	79.16	83.19
H	Total	ACA	75.30	83.19	76.70	79.06

Tables 7 and 8. This fact indicates the importance of the fusion in fault diagnosis.

6.2 Discussions

The GA-LS-WSVM method chosen from the SVM family is compared with two methods from the neural network family, i.e. Back-Propagation Neural Networks (BPNN) and GA-Recurrent NN (GARNN). The specific performance of these methods is presented in Table 10. On the other hand, the performance of these two methods using the decision level fusion with DSER is presented in Table 11. Specific performances of BPNN and GA-RNN in training and testing phase are for BPNN 74.30 and 68.60, for GARNN 74.03 and 82.65, respectively. This analysis shows that the performance of GA-RNN is better than BP-NN. The ACA performances of BPNN and GA-RNN with DSER, for training and testing phases are respectively 75.30 and 83.19for DSER-BPNN, and 76.70 and 79.06 for DSER-GARNN. Moreover, Comparing Tables 7-11, it is concluded that the performance of SVM is better than NN and fusion strategy improves the performance of

fault diagnosis. Better performance of SVM results from overcoming the issues of NN, i.e. its convergence speed is high, does not experience over-fitting and is not trapped in local extrema. Other cases that are highlighted in this paper are:

1. Using the EMD signal processing method makes it possible that at all operating points of the WT, fault diagnosis is conducted. The function EMD decomposes the signal components in a better way and performs the feature extraction in both time and frequency domains.
2. Using a method for selecting the best subset of features reduces the extra features that cause an increase in the input vector.
3. A proper choice of LS-SVM parameters influences its performance. The use of GA improves fairly the kernel and the SVM parameters.
4. The use of wavelet function instead of RBF leads to better performance of the SVM.
5. Using a hybrid fusion in feature and decision levels leads to higher classification precision for ITSC fault diagnosis.
6. Using the electrical signals existing in control system eliminates the need for extra sensors and data collection equipment (for collecting signals such as vibration and acoustic signals) with the purpose of condition monitoring and fault detection.

7 Conclusions

In this paper, a novel structure based on two-level fusion (FLF and DLF) for diagnosing ITSC fault in a DFIG-based WT, has been presented. In this paper, the repair and maintenance structure on the basis of fusion-based monitoring has been used. In the first layer, seven signals including three signals of rotor phases, three signals of stator phases and a reactive power signal are used. In the second layer and at the feature extraction level, the EMD function is used which is applicable in no stationary conditions, and 80 features were extracted from each signal. Feature matrices are very versatile and include different operation ranges of the system along with the time and frequency domain characteristics. In the third layer, features are fused at the feature level. This part includes three steps namely forming the feature matrix, choosing the best feature matrix, and initial fusion and decision-making by the SVM algorithm. In order to improve the efficiency of this algorithm, three changes were made in the conventional structure of the algorithm, and finally, GA-LS-WSVM is used. The fourth or decision layer comprises two steps. Initially, the decision profile is formed by three GA-LS-WSVM outputs, and then, based on DS theory, the final decision for the fault is made. The accuracy of this method for diagnosing the ITSC fault in different levels is tested. This method is tested for both stationary and transient conditions and showed an excellent performance. The experimental have proven very well that using fusion structure increases the reliability of fault diagnosis, and based on versatile data, the percentage of false diagnosis can be reduced. The structure proposed in this paper is used to condition monitoring of WTs with other faults in electrical and mechanical subsystems, too. The proposed method is robust against changes in the load and performance parameters.

Intellectual Property

The authors confirm that they have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property.

Acknowledgment

This research was supported by MAPNA Electric & Control, Engineering & Manufacturing Co. (MECO). We thank our colleagues from MECO Control & Condition Monitoring Center who provided insight and expertise that greatly assisted the research, although they may not agree with all of the conclusions of this paper.

CRediT Authorship Contribution Statement

M. Kamarzarrin: Idea & conceptualization, Research & investigation, Data curation, Analysis, Software and simulation, Supervision, Verification, Original draft preparation, Revise & editing. **M. H. Refan:** Idea & conceptualization, Funding acquisition, Project administration, Supervision, Verification, Original draft preparation, Revise & editing. **P. Amiri:** Idea & conceptualization, Project administration, Supervision, Verification, Original draft preparation, Revise & editing. **A. Dameshghi:** Idea & conceptualization, Research & investigation, Data curation, Analysis, Software and simulation, Supervision, verification, Original draft preparation, Revise & editing.

Declaration of Competing Interest

The authors hereby confirm that the submitted manuscript is an original work and has not been published so far, is not under consideration for publication by any other journal and will not be submitted to any other journal until the decision will be made by this journal. All authors have approved the manuscript and agree with its submission to “Iranian Journal of Electrical and Electronic Engineering”.

References

[1] Zh. Q. Wu and J. P. Xie, “Design of adaptive robust guaranteed cost controller for wind power generator,” *International Journal of Automation and Computing*, Vol. 10, No. 2, pp. 111–117, 2013.

[2] A. Dameshghi and M. H. Refan, "A new strategy for short-term power-curve prediction of wind turbine based on PSO-LS-WSVM," *Iranian Journal of Electrical and Electronic Engineering*, Vol. 14, No. 4, pp. 392–403, 2018.

[3] P. Cross and X. Ma, "Model-based and fuzzy logic approaches to condition monitoring of operational wind turbines," *International Journal of Automation and Computing*, Vol. 12, No. 1, pp. 25–34, 2015.

[4] N. Laouti, S. Othman, M. Alamir, and N. Sheibat-Othman, "Combination of model-based observer and support vector machines for fault detection of wind turbines," *International Journal of Automation and Computing*, Vol. 11, No. 3, pp. 274–287, 2014.

[5] S. Djurovic, C. J. Crabtree, P. J. Tavner, and A. C. Smith, "Condition monitoring of WT induction generators with rotor electrical asymmetry," *IET Renewable Power Generation*, Vol. 6, No, 4, pp. 207–216, 2012.

[6] M. Yousefi kia, M. Khedri, H. R. Najafi, and M. A. Shamsi Nejad, "Hybrid modelling of doubly fed induction generators with inter-turn stator fault and its detection method using wavelet analysis," *IET Generation, Transmission & Distribution*, Vol. 7, No. 9, pp. 982–990, 2013.

[7] R. Roshanfekr and A. Jalilian, "Analysis of rotor and stator winding inter-turn faults in WRIM using simulated MEC model and experimental results," *Electric Power Systems Research*, Vol. 119, pp. 418–424, 2015.

[8] T. A. Kawady, A. A. Afify, A. M. Oshiba, and A. I. Taalab, "Modeling and experimental investigation of stator winding faults in induction motors," *Electric Power Components and Systems*, Vol. 37, No. 6, pp. 599–611, 2009.

[9] R. Roshanfekr and A. Jalilian, "Wavelet-based index to discriminate between minor inter-turn short-circuit and resistive asymmetrical faults in stator windings of doubly fed induction generators: a simulation study," *IET Generation, Transmission & Distribution*, Vol. 10, No. 2, pp. 1–8, 2016.

[10] P. V. Rodríguez and A. Arkkio, "Detection of stator winding fault in induction motor using fuzzy logic," *Applied Soft Computing*, Vol. 8, No. 2, pp. 1112–1120, 2008.

[11] R. Sharifi and M. Ebrahimi, "Detection of stator winding faults in induction motors using three-phase current monitoring," *ISA Transactions*, Vol. 50, No. 1, pp. 14–20, 2011.

[12] A. Bechkaouia, A. Ameurb, S. Bourasc, and A. Hadjadjd, "Open-circuit and inter-turn short-circuit detection in PMSG for WT applications using fuzzy logic," *Energy Procedia*, Vol. 74, pp. 1323–1336, 2015.

[13] M. S. Ballal, Z. J. Khan, H .M. Suryawanshi, and R. L. Sonolikar, "Adaptive neural fuzzy inference system for the detection of inter-turn insulation and bearing wear faults in induction motor," *IEEE Transactions on Industrial Electronics*, Vol. 54, No. 1, pp. 250–258, 2007.

[14] A. Djerdira, S. S. Moosavia, Y. Ait-Amiratb, and D. A. Khaburica, "ANN based fault diagnosis of permanent magnet synchronous motor under stator winding shorted turn," *Electric Power Systems Research*, Vol. 125, pp. 67–82, 2015.

[15] R. M. Tallam, S. B. Lee, G. C. Stone, G. B. Kliman, J. Yoo, T. G. Habetler, and R. G. Harley, "A survey of methods for detection of stator-related faults in induction machines," *IEEE Transactions on Industry Applications*, Vol. 43, No. 4, pp. 920–933, 2007.

[16] I. Albizu, A. Tapia, J. R. Saenz, A. J. Mazon, and I. Zamora, "Online stator winding fault diagnosis in induction generators for renewable generation," in *Proceedings of the 12th IEEE Mediterranean Electrotechnical Conference (IEEE Cat. No. 04CH37521)*, Vol. 3, pp. 1017–1020, 2004.

[17] L. H. Hansen, L. Helle, F. Blaabjerg, E. Ritche, S. Munk-Nielsen, H. Bindner, P. Sorensen, and B. Bak-Jensen, "Conceptual survey of generators and power electronics for WTs," *Risø National Laboratory*, Roskilde, Denmark, Vol. 1205, pp. 71–73, 2001.

[18] H. Douglas, P. Pillay, and P. Barendse, "The detection of interturn stator faults in doubly-fed induction generators," in *Fortieth IAS Annual Meeting. Conference Record of the 2005 Industry Applications Conference*, Vol. 2, pp. 1097–1102, 2005.

[19] P. S. Barendse and P. Pillay, "A new algorithm for the detection of inter-turn stator faults in doubly-fed wind generators," in *32nd Annual Conference on IEEE Industrial Electronics (IECON 2006)*, Paris, France, pp. 1235–1240, 2006.

[20] Q. Lu, T. Breikin, and H. Wang, "Modelling and fault diagnosis of stator inter-turn short circuit in doubly-fed induction generators," in *IFAC Proceedings*, Milan, Italy, Vol. 44, No. 1, pp. 1013–1018, 2011.

[21] G. Stojcic and K. Pasanbegovic, "Detecting faults in doubly fed induction generator by rotor side transient current measurement," *IEEE Transactions on Industry Applications*, Vol. 50, No. 5, pp. 3494–3502, 2014.

[22] M. B. Abadi, S. M. A. Cruz, and A. P. Gonçalves, "Inter-turn fault detection in doubly-fed induction generators for WT applications using the stator reactive power analysis," in *IET 3rd Renewable Power Generation Conference (RPG 2014)*, pp. 1–6, 2014.

[23] S. Kanerva, S. Seman, and A. Arkkio, "Inductance model for coupling finite element analysis with circuit simulation," *IEEE Transactions on Magnetics*, Vol. 41, No. 4, pp. 1620–1623, 2005.

[24] Q. Lu, T. Breikin, and H. Wang, "Modelling and fault diagnosis of stator inter-turn short circuit in doubly fed induction generators," *IFAC Proceedings*, Vol. 44, No. 1, pp. 1013–1018, 2011.

[25] Y. Chen, L. Wang, Y. Zhao, and T. Tanaka, "FEM simulation and analysis on stator winding inter-turn fault in DFIG," in *IEEE 11th International Conference on the Properties and Applications of Dielectric Materials (ICPADM)*, 2015.

[26] S. Toma, L. Capocchi, and G. A. Capolino, "Wound rotor induction generator inter-turn short-circuits diagnosis using a new digital neural network," *IEEE Transactions on Industrial Electronics*, Vol. 60, No. 9, pp. 4043–4052, 2013.

[27] G. Y. Sizov, C. C. Yeh, and N. A. O. Demerdash, "Magnetic equivalent circuit modeling of induction machines under stator and rotor fault conditions," *IEEE International Electric Machines and Drives Conference*, pp. 119–124, 2009.

[28] D. Shah, S. Nandi, and P. Neti, "Stator-interturn-fault detection of doubly fed induction generators using rotor-current and search-coil-voltage signature analysis," *IEEE Transactions on Industry Applications*, Vol. 45, No. 5, pp. 1831–1842, 2009.

[29] J. Zafar and J. Gyselinck, "CUSUM based fault detection of stator winding short circuits in doubly-fed induction generator based wind energy conversion systems," in *International Conference on Renewable Energies and Power Quality (ICREPQ'10)*, Granada, Spain, 2010.

[30] S. Djurovic, S. Williamson, and A. Renfrew, "Dynamic model for doubly-fed induction generators with unbalanced excitation, both with and without winding faults," *IET Electric Power Applications*, Vol. 3, No. 3, pp. 171–177, 2009.

[31] R. Roshanfekr and A. Jalilian, "Wavelet-based index to discriminate between minor inter-turn short-circuit and resistive asymmetrical faults in stator windings of doubly fed induction generators: A simulation study," *IET Generation, Transmission & Distribution*, Vol. 10, No. 2, pp. 1–8, 2016.

[32] F. Jawad, H. Nejadi-Koti, and A. H. Exiri, "Inductance-based inter-turn fault detection in permanent magnet synchronous machine using magnetic equivalent circuit model," *Electric Power Components and Systems*, Vol. 45, No. 9, pp. 1016–1030, 2017.

[33] M. S. Ballal, H. M. Suryawanshi, and K. M. Mahesh, "Stator winding inter-turn insulation fault detection in induction motors by symmetrical components method," *Electric Power Components and Systems*, Vol. 36, No. 7, pp. 741–753, 2008.

[34] N. Gang and Y. Bo-Suk, "Intelligent condition monitoring and prognostics system based on data-fusion strategy," *Expert Systems with Applications*, Vol. 37, No. 12, pp. 8831–8840, 2010.

[35] H, B, Mitchell, *Data fusion: Concepts and ideas*. Springer Science & Business Media, 2012.

[36] X. Ye and L. Osadciw, "Applying swarm intelligence and Bayesian inference for WT SCADA-based condition monitoring and prognostics," in *Annual Conference of the Prognostics and Health Management Society*, Vol. 18, 2014.

[37] Y. Dong, L. Hui, H. Yaogang, Z. Jie, X. Hongwei, and L. Yongsen, "Vibration condition monitoring system for WT bearings based on noise suppression with multi-point data fusion," *Renewable Energy*, Vol. 92, pp. 104–116, 2016.

[38] A. Dameshghi and M. H. Refan, "Wind turbine gearbox condition monitoring and fault diagnosis based on multi-sensor information fusion of SCADA and DSER-PSO-WRVM method," *International Journal of Modelling and Simulation*, Vol. 39, No. 1, pp. 48–72, 2018.

[39] M. Khazaee, H. Ahmadi, M. Omid, and A. Moosavian, "Feature-level fusion based on wavelet transform and Artificial Neural Network for fault diagnosis of planetary gearbox using acoustic and vibration signals," *Insight-Non-Destructive Testing and Condition Monitoring*, Vol. 55, No. 6, pp. 6323–330, 2013.

[40] Z. Liu, W. Guo, Z. Tang, and Y. Chen, "Multi-sensor data fusion using a relevance vector machine based on an ant colony for gearbox fault detection," *Sensors*, Vol. 15, No. 9, pp. 21857–21875, 2015.

[41] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, Vol. 454, pp. 903–995, 1998.

[42] J. A. Rosero, L. Romeral, J. A. Ortega, and E. Rosero, "Short-circuit detection by means of empirical mode decomposition and Wigner–Ville distribution for PMSM running under dynamic condition," *IEEE Transactions on Industrial Electronics*, Vol. 56, No. 11, pp. 4534–4547, 2009.

[43] B. S. Yang, T. Han, and J. L. An, "ART-KOHONEN neural network for fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, Vol.18, pp. 645–657, 2004.

[44] Y. G. Lei, Z. J. He, Y. Y. Zi, and X. F. Chen, "New clustering algorithm-based fault diagnosis using compensation distance evaluation technique," *Mechanical Systems and Signal Processing*, Vol. 22, No. 2, pp. 419–435, 2008.

M. Kamarzarrin was born in 1991 and is receiving currently her B.Sc. degrees (with the highest honors) in Electronic Engineering from Shahid Rajaee Teacher Training University (SRTTU), Tehran, Iran, in 2013. She received her M.Sc. in Control Engineering from Shahid Beheshti University, Tehran, Iran in 2015. Now she is Ph.D. Student of Electronic Engineering in Shahid Rajaee Teacher Training University. Her research interests include GPS, wind turbine, fault detection & tolerant control, adaptive control, wireless communications and networking with a focus on cognitive radios, analog electronics, and Boolean function. Now she is a wind turbine process expert in MAPNA Electric & Control, Engineering & Manufacturing Co. (MECO).

M. H. Refan received his B.Sc. in Electronics Engineering from Iran University of Science and Technology, Tehran, Iran in 1972. After 12 years of working and experience in the industry, he started studying again in 1989 and received his M.Sc. and Ph.D. in the same field and the same University in 1992 and 1999 respectively. He is currently Professor of Electrical and Computer Engineering Faculty, Shahid Rajaee Teacher Training University, Tehran, Iran. He is the author of about 50 scientific publications in journals and international conferences. His research interests include GPS, DCS, and automation system, wind turbine, fault detection & tolerant control, and adaptive control.

P. Amiri was born in 1970. He received the B.Sc. degree from the University of Mazandaran in 1994, M.Sc. from Khajeh Nasir Toosi University (KNTU), Tehran, Iran, in 1997, and his Ph.D. from Tarbiat Modares University (TMU), Tehran, Iran, in 2010, all degrees in Electrical Engineering (Electronics). His main research interest includes electronic circuit design in industries. His primary research interest is in RF and power electronic circuits, with focus on high efficient and high linear power circuit design. He is currently an Associate Professor of Electrical and Computer Engineering Faculty, Shahid Rajaee Teacher Training University, Tehran, Iran.

A. Dameshghi was born in1986 and received his B.Sc., M.Sc. Ph.D. degrees in Electronic Engineering from the Department of Electrical Engineering, of Electrical and Computer Engineering, Shahid Rajaee Teacher Training University (SRTTU), Tehran, Iran, in 2011, 2013, and 2020, respectively. His research interests include Boolean Function, GPS, wind turbine, fault detection & tolerant control, adaptive control, electric and hybrid vehicle, and now he is the manager of EV & Infrastructure Development Center of MAPNA Electric & Control, Engineering & Manufacturing Co. (MECO).