3rd International Conference on Automation, Control, Engineering and Computer Science (ACECS'16)
Proceedings of Engineering & Technology (PET)
pp. 675-681

WANN fault detection and isolation approach design for a FAST wind turbine benchmark model

T. Bakir and B. Boussaid
Gabes University Tunisia
Emails: tahani.bakirt@gmail.com
dr.boumedyen.boussaid@ieee.org

P. F. Odgaard
Aalborg University
Denmark
Email: odgaard@ieee.org

M. N. Abdelkrim, C. Aubrun
Gabes University,Lorraine University
Emails: naceur.abdelkrim@enig.rnu.tn
christophe.aubrun@univ-lorraine.fr

***Abstract*—A design of a wavelet-based artificial neural network (WANN) classifier for a nonlinear wind turbine diagnosis system is proposed in this paper. This classifier is dedicated to evaluate residue signals and isolate faults. The WANN algorithm is implemented and tested under three different cases). First, the Discrete Wavelet Transform (DWT) with the Multi-Resolution Analysis (MRA) is applied to decompose signals of residues at resolution levels of the components of the residue signals. Approximation and details at different resolution levels are used for the artificial neural network (ANN). Second, the neural network classifier these details to identify the residue type according to three cases. The classifier has been tested under different cases events such as with faulty free, fault 1 and fault 2 scenarios. The simulation results illustrated that the proposed classifier is able to recognize and classify signals of residues efficiently and can achieve high accuracy rate under various test cases.**

I. Introduction

In recent years, the developments of many monitoring and supervision systems are being developed in order to ensure the success of planned operations and recognize the behaviorial problems in the wind turbine process. Among other functions, detection, diagnose and eliminating faults are crucial tasks of these systems ensuring the satisfaction of the performance's process operations specifications. In addition, the information supplied by a monitoring system should not only notify the system operator about what is going on, but also advice him to assume corrective actions in order to cure the problem [13]. As a result, three basic aims will be reached: the unavailing time will be minimized, the system protection will be ameliorated and the operational costs will be decreased. The process monitoring can be arranged as sequence of four states : fault detection, fault isolation, fault diagnosis and fault recovery.

Often, the operator cannot observe the automated changes from one state to another which is transparent, displaying only the crucial information to take appropriate action. As a remedy of this problem, a number of time-frequency domain techniques have been proposed including the short-time Fourier transform (STFT), the Wigner-Ville distribution (WVD), and the wavelet transform (WT) have been used [1], [10], [19].

Wavelet theory has turned into one of the signal processing tool and the emerging and fast-evolving mathematical for its many distinct benefits [17], [18]. Wavelet analysis affords a different view of data than those presented by traditional techniques, like discontinuities in higher derivatives, trends and breakdown points. Moreover, because of aspects of data that other signal analysis techniques miss, wavelet analysis affords without appreciable degradation a denoised signal. The WT is used to represent all possible types of transients in signals generated by faults in a wind turbine system.

Newly, applications of the wavelet transform (WT) and artificial neural network (ANN) in diagnosis fields [21]can be found in several studies that refer primarily on the signal processing and classification in different area [20], [21]. Recent advances in the field of neural networks have made them interesting for analyzing signals. So that, they have been successfully used in a diversity of diagnosis applications in wind turbine system [17], [19]. ANNs not only model the signal, but also make a decision as to the class of signal.

This paper presents an algorithm for classification of signals based on a combine wavelet transformation (WT) and ANN classifying techniques. Discrete Wavelet Transform (DWT) with the Multi-Resolution Analysis (MRA) is applied to decompose residue signal in approximation and details. The neural network classifies these details to identify faults.

The paper is structured as follow. The second section is dedicated for the description of the benchmark model used in our work. The methodology process is presented in Section II of this paper. Simulation results of the proposed algorithm with wavelet decompose and ANN classifications are given in Section III. Conclusions are given in Section IV.

II. Methology Description

A. Wavelet transform and multi resolution analysis

The wavelet transform (WT) introduces a useful representation of a function in the time-frequency domain in order to analyze signal structures of different sizes [14], [15], [16]. Basically, a wavelet is a function can be defined as the projection on the basis of wavelet functions . The Continuous Wavelet Transformation (CWT) of a signal $x(t)$ can be expressed as:

$$T_o(a,b) = \int_{-\infty}^{+\infty} x(t)\,\psi_{a,b}(t)dt \tag{1}$$

with

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{2}$$

The $\psi_{a,b}$ functions are obtained from the dilating and translating of the mother wavelet function $\psi(t)$. The dilating

ISSN: 2356-5608

and translating coefficients, a, $b \in \Re$, $a \neq 0$, respectively. The scale parameter a determines the oscillatory frequency and the length of the wavelet, and the translation parameter b determines its shifting position. $\psi(t)$ must have a zero mean denoted as:

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{3}$$

A real wavelet transform is complete and conserves energy as long as it satisfies a weak admissibility condition. this condition is given in the next equation whether $\psi(t) \in \pounds^2$:

$$\int_{\Re} \frac{\left|\hat{\psi}(t)\right|}{|\omega|} d\omega \prec \infty \tag{4}$$

If this condition is satisfied, we can analyze then reconstruct the signal without loss of information [14].

The WT is designed to address the problem of nonstationary signals. It involves representing a time function in terms of simple, fixed building blocks, termed wavelets. These building blocks are actually a family of functions which are derived from a single generating function called the mother wavelet by translation and dilation operations. The main advantage of the WT is that it has a varying window size, being broad at low frequencies and narrow at high frequencies, thus leading to an optimal timefrequency resolution in all frequency ranges. Furthermore, owing to the fact that windows are adapted to the transients of each scale, wavelets lack of the requirement of stationarity. The property of time and frequency localization is known as compact support and is one of the most attractive features of the WT. The WT of a signal is the decomposition of the signal over a set of functions obtained after dilatation and translation of an analyzing wavelet [14].

In engineering areas, the application of WT usually requires the discrete WT (DWT)[10], [19], [17]. The DWT is defined by using discrete values of the scaling parameter a and the translation parameter b . For better explain, set $a = a_0^m$ and $b = nb_0a_0^m$ then we get $\psi_{m,n}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0)$, where m, $n \in Z$. In general cases, we can choose $a_0 = 2$ and $b_0 = 1$. This choice will present a dyadic-orthonormal WT and define the basis for multi-resolution analysis (MRA). In MRA, any time series $x(t)$ can be completely decomposed in terms of approximations, provided by scaling functions $\phi_m(t)$ (also called father wavelet) and the details, provided by the waveletsy $\psi_m(t)$. The decomposition procedure starts by passing a signal through these filters. The scaling function is associated with the low-pass filters (LPF), and the wavelet function is associated with the high-pass filters (HPF). The decomposition procedure starts by passing a signal through these filters. The coefficients A_j and D_j are computed using the tree decomposition algorithm allowing storing low frequency information of the signal as well as the discontinuities. In fig.2 HBF, LBF, represent the decomposition filters and $\downarrow 2$ denotes a down sampling by a factor of 2. Thus, we can conclude that A_1 being the approximate version of the original signal, LBF behaves as a low pass filter. If D_1 contains only high frequency components of signal $x(t)$, then HBF behaves as a high pass filter. Finally, the signal is decomposed at the expected level.

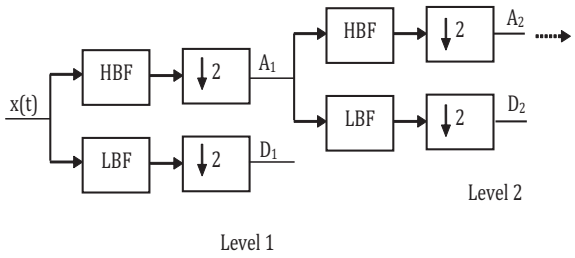

Fig. 1. MRA principal

B. ANN classifier

1) ANN classifier: An ANN is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. Through learning process, an ANN can be configured for a specific application, such as pattern recognition or data classification [20], [23], [22].

The Artificial Neural Networks have parallel structures, massively distributed, contained neuron which are simple processing units. These structures draw an analogy of the human brain due to its ability to acquire knowledge from the environment[23], [24].

Neural Network application consists of two steps: Training step and testing step. The network is trained using input data, and after it is tested. This learning take place through an adjustment of the connection weights, or synaptic weights, which exists between neurons. These connections keep the information obtained by the network.

In literature [20], [23], [22], there are various neural network architectures, such as radial basis function networks, Kohonen networks, support vector machine and so many others.

2) Multilayer Perceptron (MLP): A Multilayer Perceptron (MLP) is a feed forward neural network model. It maps sets of input data onto a set of appropriate output. Using three or more layers of neurons (nodes) with nonlinear activation functions, the MLP modify the standard linear perceptron. Thence, it is capable to distinguish data that is nonlinearly separable which makes it more powerful than the perceptron. In figure (2),the basic structure of MLP neural network is presented. In one layer, each node is connected with a certain weight. This connection is made for every other node in the next layer. The number of ANN is trained by using different algorithms like Levenberg- Marquardt algorithm (trainlm), Resilient back propagation algorithm (trainrp), Scaled conjugate gradient algorithm (trainscg), etc.

In our work, we use a Multilayer Perceptron (MLP). This choice is due to its simplicity and applicability. Levenberg-Marquardt (LMA) is the training algorithm used which is

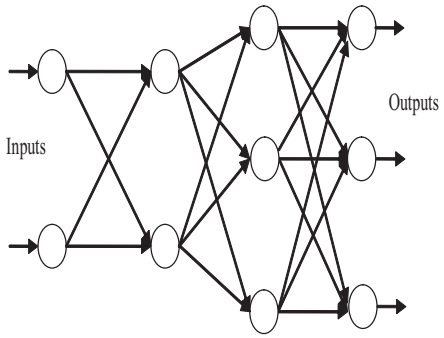

Fig. 2. A Multilayer Perceptron Structure

available in mathematical software Matlab.

3) Backpropagation Algorithm: For the training procedure, an adaptive backpropagation (BP) algorithm is used in two phases [23]. In the first phase, the inputs are introduced to the network, which propagate forward to generate the output for each neuron, $y_j(t)$, in the output layer. $y_j = f(z_j)$ determines the activity of each neuron, where $f(z) = 1/(1 + e^{-z})$ the sigmoid activation function and α and β are constants. Then the error signal is defined as fellow:

$$e_j(t) = d_j(t) - y_j(t) \tag{5}$$

$y_j(t)$ is network output, $d_j(t)$ is the desired output for neuron j at iteration t. The BP algorithm changes the weight vector, ω, of NN so as to minimize the error function, ζ, defined by:

$$\varsigma(t) = \frac{1}{2}\sum_{C\in j} e_j^2(t) \tag{6}$$

where the C includes all neurons in the output layer. The correction $\Delta\omega_{ei}$ applied to ω_{ei} is defined by:

$$\Delta\omega_{ei} = -\beta\frac{\partial\zeta}{\partial\omega_{ei}} \tag{7}$$

Training of the NN is based on an adaptive algorithm with the parameter β changing. If in (7), $\partial\zeta\backslash\partial\omega_{ei} = 0$, a minimum has been reached.
The prediction accuracy of the MLP network is based on mean square error (MSE) which is given by equation (8)

$$MSE = \frac{1}{M}\sum_{i=1}^{M}\left(d^{(i)} - y^{(i)}\right)^2 \tag{8}$$

$y^{(i)}$ is network output, $d^{(i)}$ is desired output and M is the number of training patterns (input-output pairs). For each input pair, the output of the network $y^{(i)}$ is compared with the desired output $d^{(i)}$ by computing the error given by equation (9):

$$e = d^{(i)} - y^{(i)} \tag{9}$$

III. Benchmark Model descrption

A. FAST Model descrption

The National Renewable Energy Laboratory (NREL) and its academic and industry partners have created aeroelastic simulators for horizontal-axis wind turbines (HAWTs). The NREL offshore 5-MW [1] baseline wind turbine has two or three blades [8],[9]. A generator fully coupled to a converter is used to convert the mechanical energy to electrical energy. The energy conversion from wind energy to mechanical energy is controlled by two ways. The first way is by changing the aerodynamics of the turbine by pitching the blades while the second one is by controlling the rotational speed of the turbine relative to the wind speed [3], [5]. The benchmark model for the wind turbine in this work is the FAST (Fatigue, Aerodynamics, Structures, and Turbulence) model.The benchmark model is detailed in [3]. It is a comprehensive aeroelastic simulator capable of predicting both the extreme and fatigue loads of two and three-bladed horizontal-axis wind turbines [3], [4]. It has five flexible bodies: tower, three blades, and drive shaft. Table I presents the additional gross properties for the NREL 5-MW baseline wind turbine model.

TABLE I
Gross Properties Chosen for the NREL 5-MW Baseline Wind Turbine

Rating	5 MW
Rotor Orientation, Configuration	Upwind, 3 Blades
Control	Variable Speed, Collective Pitch
Drivetrain	High Speed, Multiple-Stage Gearbox
Rotor, Hub Diameter	126 m, 3 m
Hub Height	90 m
Cut-In, Rated Cut-Out Wind Speed	3 m/s, 11.4 m/s, 25 m/s
Cut-In, Rated Rotor Speed	6.9 rpm, 12.1 rpm
CRated Tip Speed	80 m/s

The aerodynamic of the wind turbine can be represented by, [4]:

$$\tau_r(t) = \frac{\rho\pi R^3 C_q(\lambda(t),\beta(t))v(t)^3}{2\omega_r(t)} \tag{10}$$

where C_q is the torque coefficient, ρ is the air density, R is the radius of blades, v is the wind speed and β is the pitch angle. It should be noted that the estimation of the $\tau_r(t)$ is based on measured $\beta(t)$ and $\omega_r(t)$ and an estimated $v(t)$.
λ(t) is the tip speed ratio defined as:

$$\lambda(t) = \frac{\omega_r(t)\,R}{v(t)} \tag{11}$$

The estimation of $\rho(t)$ is affected by an unknown measurement error which is caused by the uncertainty of the wind speed. We can estimate also this error by means of the error of the approach described in the next section.

[1] A single 5-MW wind turbine can supply enough energy annually to power 1,250 average American homes.

Let us consider the nonlinearity represented by the relations (10) and (11), in which the wind turbine control is exploited. The model of the drive train is presented by [4]:

$$\dot{\omega}_r(t) = \frac{1}{J}\left(\tau_r(t) - \tau_g(t)\right) \tag{12}$$

with

$$\dot{\tau}_g(t) = p_{gen}(\tau_{ref}(t) - \tau_g(t)) \tag{13}$$

where $\tau_g(t)$ is the generator torque and $\tau_{ref}(t)$ is its reference. p_g is the generator power. The wind turbine model description in the continuous-time domain is. :

$$\begin{cases} \dot{x}_c(t) = f_c(x_c(t), u(t)) \\ y(t) = x_c(t) \end{cases} \tag{14}$$

where $u(t) = [\tau_{ref}(t), \tau_r(t)]^T$ and $y(t) = x_c(t) = [\omega_r(t), \tau_g(t)]^T$. $f_c(.)$ represents the continuous-time nonlinear function that will be subsequently approximated via discrete-time fuzzy prototype from N sampled data $u(t)$ and $y(t)$, with $t = 1, 2, ..., N$. Lastly, the model parameters and the map C_p (λ, β) are chosen as given in [4] in order to represent a realistic turbine.

For the residue generation, sensor faults of the system under diagnosis are treated based on the data driven measurement for uncertain sequences $u(t)$ and $y(t)$. A model-based approach is used to estimate the outputs of the system from a data drive[1],[2],[6], .

In this work, we consider a techniques based on fuzzy logic for modeling the nonlinear processes [1], [6], [2]. The fuzzy identification and modeling is based on 'Takagi-Sugeno' (TS) model. This method consists to approximate the non linear model by local affine models. The decomposition of input-output data acquired from actual process. We consider this method of identification to identify the wind turbine non linear model which is an unknown nonlinear model based on some available input-output data [7]. The residues r expression is as fellow:

$$r(t) = \hat{y}(t) - y(t) \tag{15}$$

Where y and $\hat{y}$ are respectively the output data driven acquired from the system and its estimated given by the block of fuzzy identification. The faulty generated signal of the residue will be treated with a suitable wavelet and artificial neural network in order to detect fault.

B. Faults descreption

Figure 3 presents the major components of the benchmark model implemented on Simulink. The feedback loop uses information from sensors as input to the pitch, torque, and yaw controllers. Actuator models for the pitch drives, generator, and yaw drive are implemented within the Simulink environment. Faults shown in this figure can corrupt both actuators and sensors. As it has been described in [3], sensors faults include measurements that are offset from the true values, stuck, scaled from the true values. In the table II, faults scenarios are illustrated.

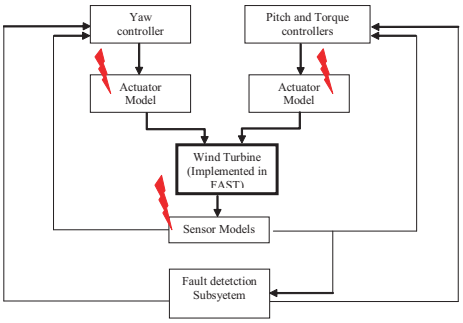

Fig. 3. Block diagram: Fault Detection and Fault Tolerant Control

TABLE II
FAULTS SCENARIOS

Fault	Type	Time(sec)
Blade root bending moment sensor	Scaling	20-45
Accelerometer	Offset	75-100
Generator speed sensor	Scaling	130-155
Pitch angle sensor	Stuck	185-210
Generator power sensor	Scaling	224-265
Low speed shaft position encoder	Bit error	225-320
Pitch actuator	Abrupt change in dynamics	370-390
Pitch actuator	Slow change in dynamics	440-465
Torque offset	offset	495-520
Yaw drive	stuck drive	550-575

IV. SIMULATION RESULTS

In our work, we consider the residue of the rotor speed angular. Using the method of generation residue signals described in [1], [2], [7]. Let consider the system 14. The output of the system is affected by faults scenarios described in table II. In our case, we consider only fault 5 from table II and fault 9 for the two scenarios faulty cases . The object of wavelet analysis is to decompose signals into several details and approximations. For the analysis of signals using DWT, two choices are crucial: the selection of appropriate wavelet and the number of decomposition levels. The number of decomposition levels is chosen based on the dominant frequency components of the signal. The level resolution is given by the next expression:

$$n = Log_2(N) - 1 \tag{16}$$

where N is the is the number of samples of the signal.

The levels are chosen such that those parts of the signal that correlate well with the frequencies necessary for classification of the signal are retained in the wavelet coefficients. In this work, Daubechies 4 (db4) is used. Its smoothing feature was adapted for detecting changes of the residue signals. Daubechies wavelets are the most popular wavelets representing foundations of wavelet signal processing, and are used in numerous applications. A detailed discussion about the

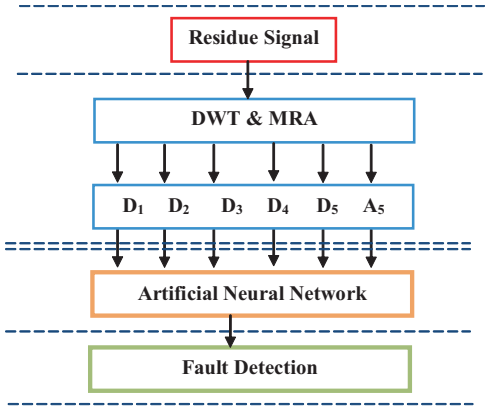

Fig. 4. Block diagram for classification of residue signals

characteristics of these wavelet functions can be found in the reference.

The wavelet decomposition is used to diversify and improve the neural network training signals. An algorithm block diagram for classification of residue signals is presented on figure (4). The algorithm structure is based on two stages: wavelet decomposition and multiresolution analysis as the first stage and classification stage the second one. The input of the second stage is a preprocessed signal.

In our case, a binary scheme of classification is used to establish the fault detection condition at the output of the classifier, namely R_N (1 0 0), R_1 (0 1 0), and R_2 defect (0 0 1) to denote three classes of the residue. These Data are used for the step of the ANN training. Only one hidden layer with different numbers of neurons in hidden layer, nh = {5,10,15,20,25} are used. The sigmoid activation function is used in the hidden and the output layer. A mean square error of $10e^{-2}$, a minimum gradient of $1e^{-10}$, and maximum number of epochs of 200 are used. The training process would stop if any one of these conditions is met. The initial weights and biases of the network are fixed randomly. The MLP neural network (MLPNN) has been implemented by using the MATLAB Neural Network Toolbox. The result of the learning process of the proposed MLP neural network and the classification MSE, are depicted in table III.

Table III demonstrates the accuracy of the classifier for varying number of hidden layer perceptron. As the number of neurons increases, the execution time increases, but the accuracy also increases.

TABLE III
ANN ARCHITECTURE AND TRAINING PARAMETERS

nh	epoch	MSE	Gradiant	Time(min)
5	60	0.202	0.0518	1.36
10	46	0.193	1.44	1.45
15	65	0.188	9.02 e^{-5}	2.54
22	49	0.186	0.00137	3.45
25	86	0.183	0.00025	9.39

After training, the wavelet- artificial neural network based

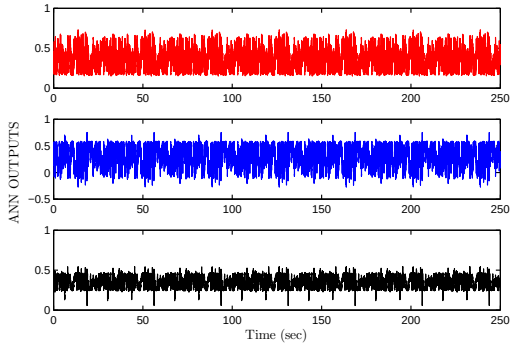

Fig. 5. ANN Outputs for residue signal with no fault

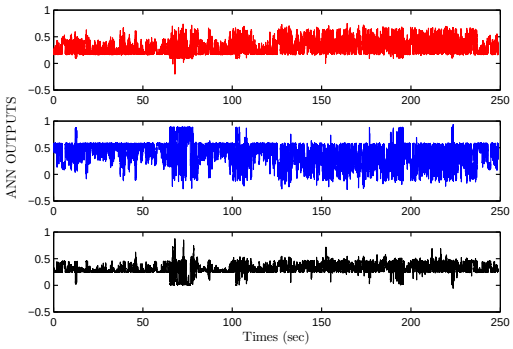

Fig. 6. ANN Outputs for residue signal with fault1

fault detector and classifier is widely tested using independent data sets consisting of fault scenarios not used previously for training the network. Fault scenario and fault location are changed for various faults in the validation/test data sets in order to investigate the main effects of these factors on the performance of the suggested algorithm.

To highlight the proposed network performance, we will try to inject three cases and see the behavior of the network. This is presented in the following scenarios. The nominal case with no fault is considered in the first scenario. Figure 5 shows the three outputs of the network which given a binary value around (1,0,0) equivalent to R_N.

Scenario 1 for fault 1 is proposed in this example. In this case, on the contrary of the training date, we change the period of the fault appearance. The fault takes place between 60 s and 80 s. Figure 6 shows the three outputs of the network which given a binary value around (0,1,0) which correspond to R_1 between the period 60 s and 80 s.

In scenario 3, faulty residue signal with fault 2 is presented. The period of the appearance of the fault is between 115 s and 125 s. Figure 7 shows the three outputs of the network which given a binary value around (0,0,1) which correspond to R_2 between the period 60 s and 80 s.

The suggested network system localize and identify the source of failures in the proposed cases. DWT/MRA and MLP architecture is useful for detecting and locating faulty in the

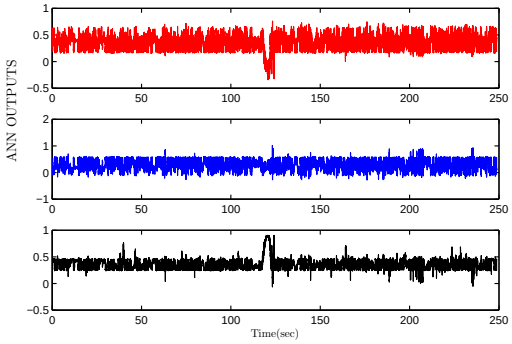

Fig. 7. ANN Outputs for residue signal with fault2

monitored wind turbine system.

V. Conclusion

In this paper, a new approach to develop classifier for identifying faults in a wind turbine benchmark model is proposed. This approach is a wavelet-artificial neural network. As residue signals, generated from a nonlinear system wind turbine, are non-stationary, the conventional method of analysis is not highly successful in diagnostic classification. In this paper, an algorithm for classification of residue signal based on DWT and ANN has been proposed. DWT with the MRA is applied to decompose residue signals at resolution levels of the components of the residue signal. The ANN classifies these generated signals to identify the fault type according to a binary output. The accuracy rates achieved by The results showed that the proposed classifier has the ability of recognizing and classifying residue signals efficiently.

References

[1] Bakir, T and Boussaid, B and Hamdaoui, R and Abdelkrim, MN and Aubrun, C Qualitative diagnosis of wind turbine system based on wavelet transform *IEEE, 15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA),*, 405–410, 2014.

[2] Bakir, T and Boussaid, B and Hamdaoui, R and Abdelkrim, MN and Aubrun, C Fault detection in wind turbine system using wavelet transform: Multi-resolution analysis *IEEE, 12th International Conference on Systems, Signals & Devices (SSD)*, 1–6, 2015.

[3] Odgaard. P.F, J. Kathryn Wind turbine fault detection and fault tolerant control : An enhanced benchmark challenge. In *Proceedings of the American Control Conference*, pages 4447-4452, 06.2013.

[4] Odgaard. P.F, S. Jakob, N. Rasmus, and D. Chris. Observer based detection of sensor faults in wind turbines. In *Proceedings of European Wind Energy Conference*, pages 1–10, 2009.

[5] Johnson, Kathryn E and Pao, Lucy Y and Balas, Mark J and Fingersh, Lee J. Control of variable-speed wind turbines: standard and adaptive techniques for maximizing energy capture. In *Control Systems, IEEE*, pages 70–81,2006. 2009.

[6] Simani.S Application of a data-driven fuzzy control design to a wind turbine benchmark model. *Advances in Fuzzy Systems*, 2012:1, 2012.

[7] Simani. S, C. Paolo, and T. Andrea. Datadriven approach for wind turbine actuator and sensor fault detection and isolation. In *Proceedings of IFAC World Congress*, pages 8301–8306, 2011.

[8] Musial, W., Butterfield, S., and Boone, A. Feasibility of Floating Platform Systems for Wind Turbines, In *the 42nd AIAA Aerospace Sciences Meeting and Exhibit*, pages pp. 57, January 2004.

[9] Jonkman, Jason Mark and Butterfield, Sandy and Musial, Walter and Scott, George Definition of a 5-MW reference wind turbine for offshore system development, In *National Renewable Energy Laboratory Golden, CO, USA*, 2009.

[10] Odgaard, Peter Fogh and Stoustrup, Jakob Wavelet based feed forward wind gust control of wind turbines, *The European Wind Energy Association*,2013.

[11] S. Lesecq, S. Gentil, I. Fagarasan Fault isolation based on wavelet transform *Journal of Control Engineering and Applied Informatics*,vol. 9:pp.51–58, 2007.

[12] T. Baoping, L. Wenyi, and S.Tao. Wind turbine fault diagnosis based on morlet wavelet transformation and wigner-ville distribution. *Renewable Energy*, 35(12):2862–2866, 2010.

[13] Y. Wenxian, PJ. Tavner, and MR. Wilkinson. Condition monitoring and fault diagnosis of a wind turbine synchronous generator drive train. *Renewable Power Generation, IET*, 3(1):1–11, 2009.

[14] Daubechies, Ingrid and others Ten lectures on wavelets *SIAM*, 1992.

[15] Wen-jing, Zhou and Yan-xia, Shen and Long, Wang Fault diagnosis for wind turbine gearbox based on wavelet analysis, *Control and Decision Conference (CCDC), 2012 24th Chinese*, pp.11638–1641, 2012.

[16] Chen, Jinglong and Pan, Jun and Li, Zipeng and Zi, Yanyang and Chen, Xuefeng Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals, *Elsevier, Renewable Energy, 2012 24th Chinese*, ,vol. 89:pp.80–92, 2016.

[17] Schmitt, Erick and Idowu, Peter and Morales, Aldo Wind turbine fault diagnosis based on Morlet wavelet transformation and Wigner-Ville distribution, *ElsevierRenewable Energy*, ,vol. 35:pp.2862–2866, 2010.

[18] Schmitt, Erick and Idowu, Peter and Morales, Aldo applications of wavelets in induction machine fault detetction *Ingeniare. Revista chilena de ingeniería*, ,vol. 18:pp.158–164, 2010.

[19] S. Lesecq, S. Gentil, I. Fagarasan* Fault isolation based on wavelet transform. *Journal of Control Engineering and Applied Informatics*, 9(3, 4):51–58, 2007.

[20] Bhaskar, Kanna and Singh, SN AWNN-assisted wind power forecasting using feed-forward neural network, *IEEE Transactions on Sustainable Energy* ,pp306–315,(2012).

[21] Catalão, JPS and Pousinho, HMI and Mendes, VMF Short-term wind power forecasting in Portugal by neural networks and wavelet transform, *Elsevier,Renewable Energy* ,pp 1245–1251,2011.

[22] egnanarayana, B Artificial neural networks, *PHI Learning Pvt. Ltd.*

[23] Dayhoff, Judith E and DeLeo, James M Artificial neural networks, Wiley Online Library *, pp 1615–1635, 2001,*

[24] Hagan, Martin T and Menhaj, Mohammad B Training feedforward networks with the Marquardt algorithm IEEE Transactions on Neural Networks*, pp 989–993, 1994,*