
Automated Fault Detection System for Wind Farms

Deepshikha Agarwal

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.68844>

Abstract

Due to the pollution and health hazards of nonrenewable resource-based energy generation systems, now focus is on the use of renewable resources. This chapter aims at providing an automated fault-detection system for increasing the robustness of offshore located wind farms. The method is based on the use of flexible threshold for calculation of the collected sample values. A fuzzy inference system (FIS) is designed for the automatic real-time fault detection system named as FIS-based fault detection system (FFDS) for offshore wind farms. The method uses the concept of combination-summation (CS) and flow-directions to determine the extent of fault occurrence in the wind farm. Based on the working conditions of the wind farm, preventive or corrective measures are suggested to the remote observer. The performance of these methods is evaluated on MATLAB.

Keywords: wind farm, wireless sensor networks, threshold, fault detection, network lifetime

1. Introduction

Wind energy is freely available everywhere in abundance. It is a renewable resource that will never get exhausted. This energy if properly utilized can lead to greener and safer energy generation compared to coal generated electricity. It is also one of the lowest priced renewable energy technologies available nowadays [1].

In 2015, energy produced in the United States was about 91% of U.S. energy consumption due to less import of petroleum [2]. Majority of energy production being due to fossil fuels, i.e., coal, petroleum, and natural gas. According to Ref. [2], natural gas contributed 32% of total generation, petroleum 28%, coal 21%, renewable energy 11%, and nuclear electric power 9%.

However, using natural gas for energy generation has several issues. First, leakage of methane during drilling and extraction of natural gas from wells and its transportation in pipelines [3]. Methane is stronger than CO₂ at trapping heat and causing global warming. Methane emissions range from 1 to 9% of total life cycle emissions. Natural gas-fired power plants contribute to acid rain and ground-level ozone, both of which can damage forests and agricultural crops [4].

The present renewable energy-based generation plants such as offshore wind farms are not entirely capable of fulfilling the future needs of the society. Due to this reason, wind-based energy generation is still not very popular and is unable to replace coal or natural gas-based energy production. The monitoring and control systems used are now obsolete and new methods are required.

The control and maintenance actions require complete human interference, and it is a time-consuming process. These challenges lead to extra cost on emergency maintenance, component screening, and physical designs.

Wind turbines consist of several components and are subject to various failures of electrical and mechanical nature [5], e.g., imbalance in electrical controls, gearbox, and yaw system. Some are more frequent and cause larger downtime of the whole system. These faults cause rotor imbalance, unbalances and harmonics in air gap flux, increase torque pulsation, and increase losses and reduction in efficiency by directly affecting the power, current, and voltage output of the generator. Therefore, monitoring of these critical components should be on the highest priority so that plant downtime can be reduced. The offshore located wind turbine generator system requires monitoring of parameters such as sea-surface temperature, wind velocity, water salinity, wave heights, and strain measurement [6, 7]. However, the monitoring of wind turbine parts has several practical difficulties, e.g., limited accessibility, large size and complex geometry of the blades, effect of environmental parameters, etc.

Several papers have discussed methods to detect faults in wind farms, e.g., gearbox fault detection using discrete wavelet transformation [8]. Similarly, high frequency vibration data collected from gearbox testing were used to gearbox fault detection in Ref. [9], which included k-means clustering algorithm. The drawbacks of this system are the assumption that the underlying process is stationary and the time factor is eliminated. Brandão et al. [10] discuss neural networks for fault forecasting of wind turbine gearbox. Badihi [11] presents protection of against the decreased power generation caused by turbine blade erosion and debris on the blades. A fault diagnosis method based on signal analysis and recognition is presented [12]. Time-frequency representations have been proposed in the literature [13–15]. These techniques have high complexity and poor resolution [16]. One approach used Hilbert transformation in a doubly fed induction generator-based wind turbine [17].

Hence, there should be some automated systems to remotely monitor these parameters and notify about faults in the system. By using wireless sensor networks (WSNs), we can ensure reliable operation of wind farm. This helps in reducing manual interference and wind farm can be completely monitored for 24 hours every day. The following sections discuss how this can be performed.

2. Flexible threshold selection scheme

In the past, the monitoring systems used constant threshold to record the data independent of time of the day or month. The constant threshold is calculated as the average of the dataset. As a result of several observations, it can be concluded that such a scheme does not give accurate results if there are changes in the scene or environment pertaining to parameters under consideration, e.g., the temperature of air during daytime is higher compared to night time. Similar variation is true during different seasons, e.g., average temperature during winter season is different from the average temperature during summer season. Hence, if constant or fixed threshold value is chosen for the entire dataset, it is likely to give unoptimized results for both the scenes. Moreover, if the chosen fixed threshold value is very high, it will result in many missed detections, and if it is very low, it will lead to many false positives.

Hence, threshold value should be selected using an appropriate scheme that allows dynamic change in the threshold value to accommodate the variations in time of data recording. This method gives better performance in terms of sensed parameters. The threshold provides a reference for finding values that are higher or lower than the threshold both of which may indicate health failures in the wind farm.

The WSN topology in wind farm consists of tower fixed nodes [18]. These are wireless sensor nodes attached to the tower nodes that can continuously sense the parameter values (samples) throughout the day and night. This information is converted into data packets that are transmitted to the sink node by taking multiple hops through the scattered sensor nodes. The sink node is located at the end of the wind farm. Every tower-fixed node is allocated a fixed local unique address called as RTN id (row-tower-node), which is transmitted as an identification of the originator of packet.

Suppose X_D is a set of samples collected by the tower-fixed sensor nodes during the day period, where

$$X_D = \{X_1, X_2, X_3, \dots, X_i, \dots, X_N\} \quad (1)$$

and Y_N is a set of samples collected by the tower-fixed sensor nodes during the night period, where

$$Y_N = \{Y_1, Y_2, Y_3, \dots, Y_i, \dots, Y_N\} \quad (2)$$

The samples collected during the night period.

The decision of choosing a new threshold for the dataset depends on the correlation between the datasets. The correlation is the measure of the similarity content between the two datasets. If the correlation of the two datasets is high, it means that the two datasets correspond to the similar time duration of the collected data and hence eliminate the need for calculating another threshold for the new dataset. Similarly, low correlation is indicative of large variations and necessitates the calculation of new thresholds for better data interpretation.

The correlation between the two datasets $R(X_D, Y_N)$ can be expressed as [19]:

$$R(X_D, Y_N) = \frac{\sum_{i=1}^N (X_i - X_m) * (Y_i - Y_m)}{\sqrt{\sum_{i=1}^N (X_i - X_m)^2 * \sum_{i=1}^N (Y_i - Y_m)^2}} \tag{3}$$

where X_i and Y_i are the values of datasets X_D and Y_N at “ i ” time instant. X_m and Y_m are the average values of the datasets, X_D , Y_N , and N is the number of samples in each dataset which should be the same for X_D and Y_N .

Figure 1 shows the scatter plot for wind speed dataset and its computed correlation coefficient. **Table 1** shows the degree of similarity between the datasets depending on the calculated correlation coefficients.

To calculate the thresholds, T_X and T_Y , the method prefers geometric mean of the datasets with “ N ” samples, instead of arithmetic mean given as below:

$$T_X = \sqrt[N]{x_1 * x_2 * \dots * x_N} \tag{4}$$

$$T_Y = \sqrt[N]{y_1 * y_2 * \dots * y_N} \tag{5}$$

We consider geometric mean because the datasets are characterized by a majority of similar

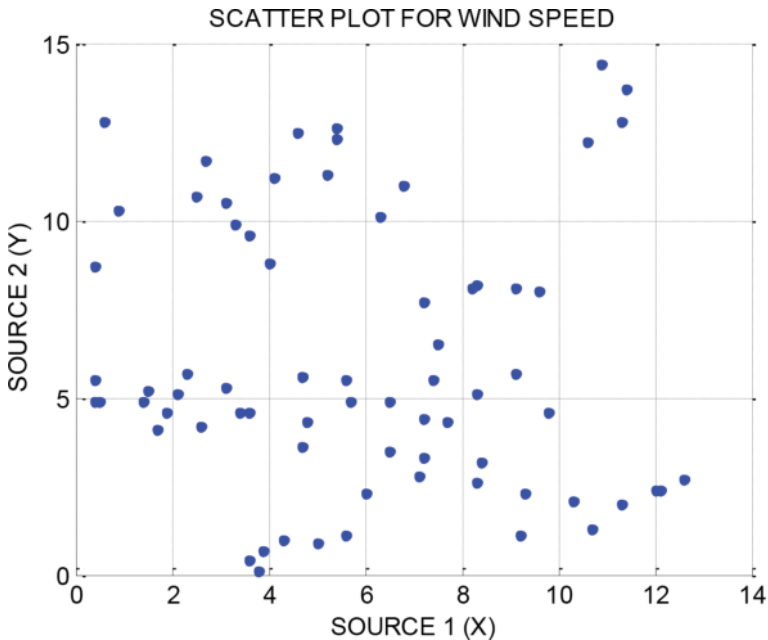


Figure 1. Scatter plots wind speed. Corr_coeff = -0.07846169 [18].

| Sl. No. | Corr_coeff. | Degree of similarity | New threshold required |
|---------|-----------------|----------------------|------------------------|
| 1 | -0.7 to +0.7 | Low | Yes |
| 2 | Less than -0.71 | High | No |
| 3 | More than +0.71 | High | No |

Table 1. Criterion for new threshold selection [18].

values and relatively very few extreme values. Being strongly influenced by few extreme observations, the threshold values calculation based on arithmetic means would fail to provide a real means of identifying the extreme values.

The use of a geometric mean normalizes the range being averaged, so that no range dominates the weighting, and a given percentage change in any of the properties has the same effect on the geometric mean. **Table 2** shows the calculated threshold values for the flexible threshold method and the mean method (MM).

Furthermore, the method requires ranging of the infinite sample values into discrete levels without changing the meaning of information using quantization. To do this, first, the distance matrices d_X and d_Y are calculated as below

$$d_X = [X_D - T_X] \tag{6}$$

$$d_Y = [Y_N - T_Y] \tag{7}$$

where these matrices represent the values of X_D and Y_N after thresholding where,

$$d_X = \{x_1', x_2', x_3', \dots, x_i', \dots, x_N'\} \tag{8}$$

$$d_Y = \{y_1', y_2', y_3', \dots, y_i', \dots, y_N'\} \tag{9}$$

Finally, the quantization is performed on the above values independently with respect to their maximum and minimum values. This can be expressed as

$$Q\alpha(\max(d) - \min(d)) \tag{10}$$

where Q is the number of quantization levels for distance matrix d .

| SRC | Variation range | Wind speed | | | | |
|------|-----------------|--------------|-------------|-------------------|-----------------|--------------------|
| | | Range values | Corr. coeff | FTS (T_{FTS}) | MM (T_{MM}) | $T_{FTS} - T_{MM}$ |
| 1(X) | (Min) | 0.4000 | -0.0784 | 10.6938 | 5.9778 | 4.7160 |
| | (Max) | 12.6000 | | | | |
| 2(Y) | (Min) | 0.1000 | | 12.6125 | | 6.6347 |
| | (Max) | 14.4000 | | | | |

Table 2. Threshold values [18].

The present scenario considers five quantization levels 0, 1, 2, 3, and 4 calculated as from Eq. (10). If the variation in the datasets and the total number of samples in it is large, the number of levels may also increase for better accuracy in fault prediction. This would lead to increase in the size of transmitted packets because the number of bits required to encode each level into binary will also increase. This would cause greater energy depletion in packet transmission, reception by the sensor nodes thus lowering the WSN network lifetime. Thus, the choice of the number of quantization levels should be able to provide accurate fault prediction without compromising the network lifetime. Each level carries a significant and distinct meaning regarding the sensed value, e.g., level “0” indicates that there is no difference between the sample value and the threshold. Similarly, levels “2,” “3,” and “4” indicate increased levels of variation. **Figure 2** depicts the above method.

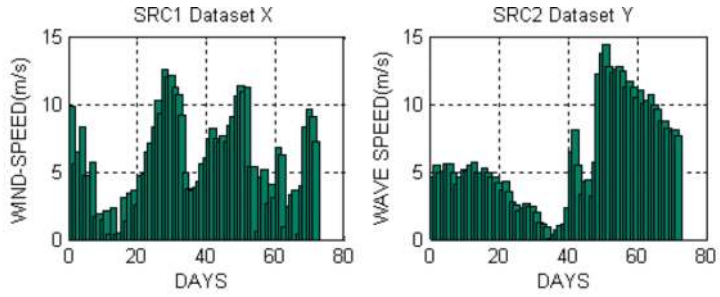
2.1. Simulation results and discussion

The flexible threshold selection (FTS) method is compared with the mean method (MM). We have considered a total of 72 samples collected during daytime and nighttime for wind speed. The sampling frequency is 1 sample per 10 minutes over a period of 12 hours daytime and 12 hours nighttime. As observed by the simulation results, the flexible threshold method gives a better performance and accurate results for parameter monitoring. **Table 2** depicts the range of collected samples and their calculated thresholds using flexible threshold selection (FTS) and mean method (MM). As observed, the datasets for source 1(X) and source 2(Y) have small variations. If both the datasets from the sources are instead, considered to be one single dataset, the variation of values is large. This causes the static threshold selected using the MM method tends to be biased toward the higher values. However, this is not the case with dynamic threshold. We can calculate different thresholds for datasets collected at different times, which will adapt with the true variations of the values known to nature. Hence, the flexible method is unbiased toward any extreme values and gives a balanced view of the data under consideration. The MM method does not consider computing new threshold every time but it remains unchanged for any dataset making it an unrealistic choice.

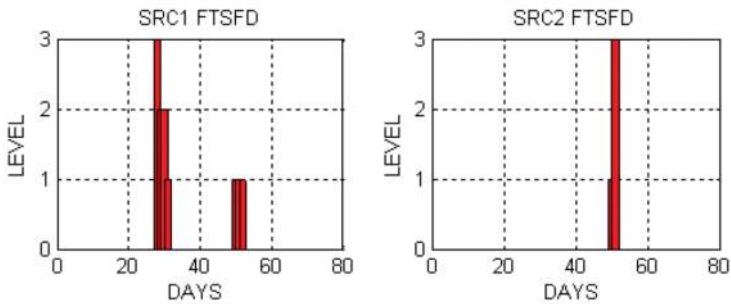
Two different thresholds for both the sources find the correlation between them by considering them individually.

It is clear from the above discussion that the choice of appropriate threshold has a large impact on the quantization levels. The MM method for threshold selection is only able to detect large variation in the values, i.e., levels “2” and “3” whereas in the FTS method the detected levels have a distributed pattern, i.e., it can detect both small and large variations. Also, the levels detected by the FTS method is consistent compared to the MM method, which provides a very accurate status of the conditions of the wind farm.

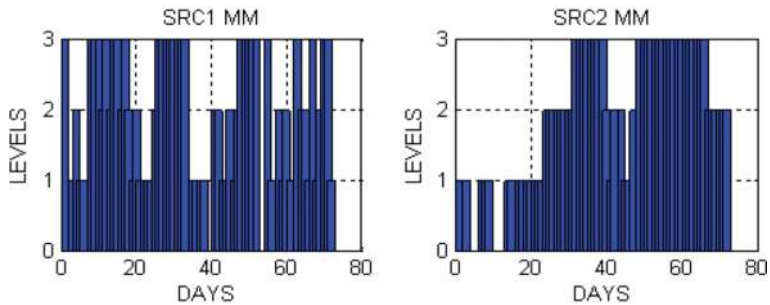
Figure 2 shows that the FTS method suggests a majority of level “0” occurrences over other levels unlike the MM method where the majority is level “2” occurrences. Thus, it can be concluded that the flexible method is unbiased toward the larger values in the datasets and



(i) Maximum-minimum variations



(ii) FTS method



(iii) MM method

Figure 2. Comparison between MM and flexible method for monitored parameter wind speed [19, 20]. (i) Collected samples of maximum and minimum wave heights, (ii) quantized levels using the FTS method, and (iii) quantized levels using the MM method.

hence provides better accuracy of monitored parameters. The graph in **Figure 2** is generated from real-time data from the Burbon-Nysted wind farm, Denmark.

3. Fault detection scheme

This system is called a fuzzy inference system (FIS)-based fault detection method (FFDS). This is an automated system, gives precise information of the health condition of the wind farm to the remote observer, and gives alarms for taking corrective or preventive measures for maintaining the reliability of the farm.

The observer needs not observe all the properties of the parameter values as a single signal, rather, the degree of similarity between the values finds the basis for choosing a new threshold. This is a simple method that helps in finding real-time data for monitoring purposes. These data when analyzed can predict all possible fault occurrences.

3.1. Automatic fault diagnosis method

The fault detection scheme uses combination-summation (CS) and flow directions (FDs) to design the FIS [9]. This aids to derive significant information from the quantized levels about fault event occurrence in the monitored data samples of offshore wind farm. The received quantized levels corresponding to monitored data samples represent the surrounding environmental conditions in offshore wind farm. The received values being fuzzy in form use FIS to provide accurate interpretation of the environmental conditions.

For determining the CS and FD, five consecutive received levels are considered in one period of time “ T ” where

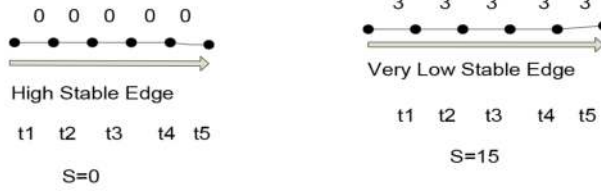
$$T = t_1 + t_2 + t_3 + t_4 + t_5 \quad (11)$$

which represents five consecutive time intervals. Depending on the permutation and the summation of the levels, fuzzy logic is used to predict fault occurrences. For example, consider the levels received at “ t ” times are $l_{t_1}, l_{t_2}, l_{t_3}, l_{t_4},$ and l_{t_5} then the summation of levels is

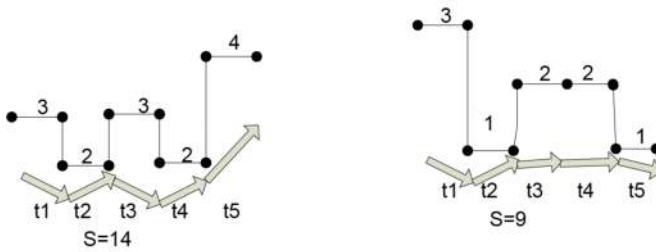
$$CS = \sum_{i=t_1}^{i=t_5} l_i = l_{t_1} + l_{t_2} + l_{t_3} + l_{t_4} + l_{t_5} \quad (12)$$

where the range of CS is [0–20]. The numeral 20 indicates constant occurrence of level 4, i.e., 44,444. The obtained levels can be either repeating or nonrepeating, e.g., 22,222, 31,224, 01,234, and 43,210 as depicted in **Figure 3**. The CS for these levels is 10, 12, 10, and 10. As observed, this alone is not sufficient for fault prediction. Fault prediction can give accurate results if the corresponding FD is also considered with the values. Here, FD means whether the received levels are in state of increasing, decreasing, remaining stable, or varying constantly. For example, if the CS is 10, it has multiple values, but if FD has raising edge, it means the combination suggests fault event occurrence in the future and calls for immediate preventive action. If the levels increase constantly, then FD is considered to be raising edge shown by arrow in upward direction (**Figure 3**).

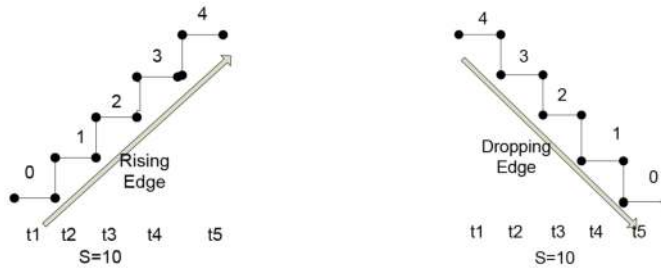
The remote observer is able to predict meaningful information from these received quantized levels based on the fuzzy-logic rules as presented in **Table 3**.



(a) Stable flow



(b) Average flow



(c) Raising/Dropping flow

Figure 3. Illustrations of repeating-level and nonrepeating-level combinations. (a) Stable flow, (b) average flow, and (c) raising/dropping flow.

In this system, the fuzzy-set “ F ” can be described [7, 8] as

$$F = \{\omega, m(\omega) | \omega \in U\} \tag{13}$$

$$U = \{0 - 3, 1\} \tag{14}$$

$$m : \omega \rightarrow [\text{Normal operation, Low risk, High risk}] \quad (15)$$

| Sl. No. | Flow directions | Combination-summations | Result (MF) |
|---------|-----------------|------------------------|----------------|
| 1 | Stable (S) | 0, 5 | Normal |
| 2 | Stable | 10 | Low risk |
| 3 | Stable | 15 | High risk |
| 4 | Stable | 20 | Very high risk |
| 5 | Rising (R) | 10 | High risk |
| 6 | Dropping (D) | 10 | Low risk |
| 7 | Average (A) | 1–4 | Normal |
| 8 | Average | 5–6 | Low risk |
| 9 | Average | 7–9 | High risk |
| 10 | Average | ≥ 10 | Very high risk |

Table 3. Fuzzy rule base.

where ω is the combination-summation and flow-directions of the received levels by the remote observer, $m(\omega)$ is the membership function for the received level and U is the universal set representing the set of all levels, as shown in **Table 3**. The membership function alerts the remote observer whenever the probability of fault occurrence becomes high (**Figure 4**). It can be formulated with risk R_i as

$$(\text{FD}_i, \text{CS}_i) \rightarrow R_i \quad (16)$$

This method is very simple to implement and efficient in enhancing the WSN lifetime.

Table 4 shows the comparison for wind speed data computed from **Figure 2**. The results confirm the belief that FFDS is able to predict accurate conditions of the wind farm. As shown, it predicts normal operation of the farm, whereas MM is only able to detect extreme values of level “0” and “3.” This leads to false alarm for corrective measures due to inaccurate calculations. Thus, it can be concluded that the FTS method is unbiased toward the larger values in the datasets and hence provides better accuracy of monitored parameters.

3.2. Simulations and discussion

Table 5 provides the details of simulation parameters used in the study. The sink node is located at the farthest point in the field. **Figure 5** shows the round in which all the nodes in the area become dead (network-lifetime).

The method FTS gives an accurate view of the parameter values in real time and the threshold selection does not indicate any biasing toward a particular value, which is confirmed from **Table 4**. Also, as observed from **Figure 5**, the network lifetime of WSN network is also increased by nearly 10 times with a packet size of 23 bits.

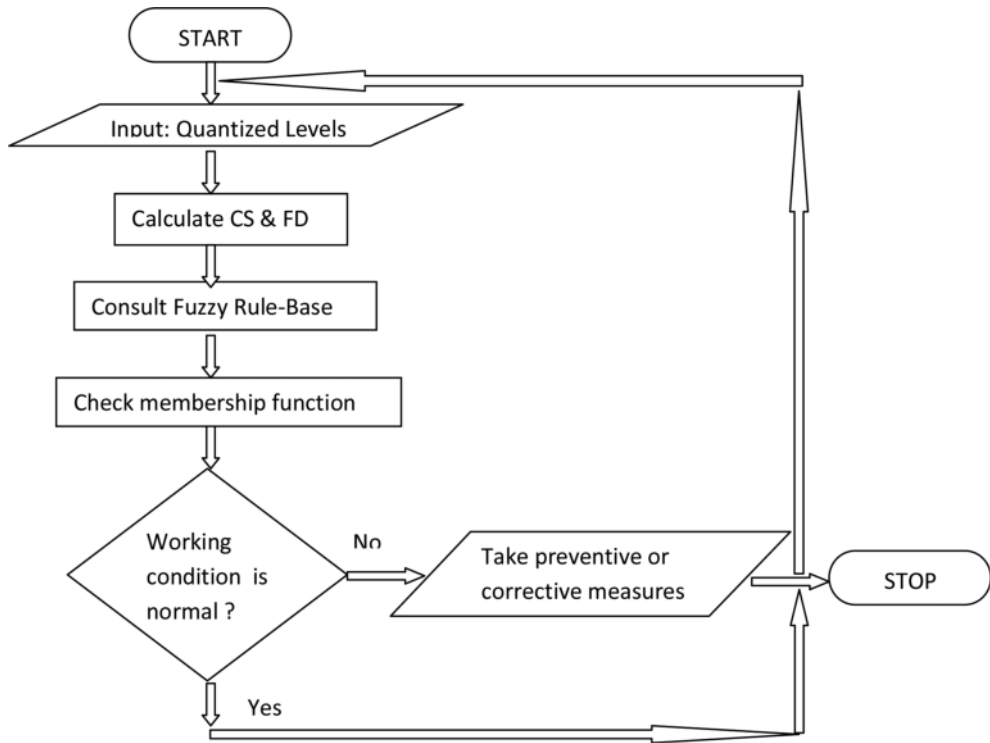


Figure 4. Flowchart to depict working of fuzzy inference system.

| Levels | Number of quantized levels | | Number of consistent quantized levels | | |
|--------|----------------------------|----|---------------------------------------|------------------|--|
| | FTS | MM | FTS | prediction | MM Prediction |
| "0" | 136 | 19 | 32 | Normal operation | 1 Fault has occurred (corrective measures) |
| "1" | 5 | 37 | 0 | | 3 |
| "2" | 2 | 37 | 0 | | 3 |
| "3" | 1 | 51 | 0 | | 9 |

Table 4. Wind speed [19].

| Sl. No. | Parameter | Value |
|---------|------------------------------|---|
| 1. | Area size | 1000 m × 1000 m, 2000 m × 2500 m, 300 m × 4500 m, 4000 m × 6000 m |
| 2. | Total number of nodes | 416–1216 |
| 3. | Total number of fixed nodes | 216 |
| 4. | Total number scattered nodes | 200–1000 |
| 5. | Total number of turbines | 72 |

Table 5. Simulation parameters.

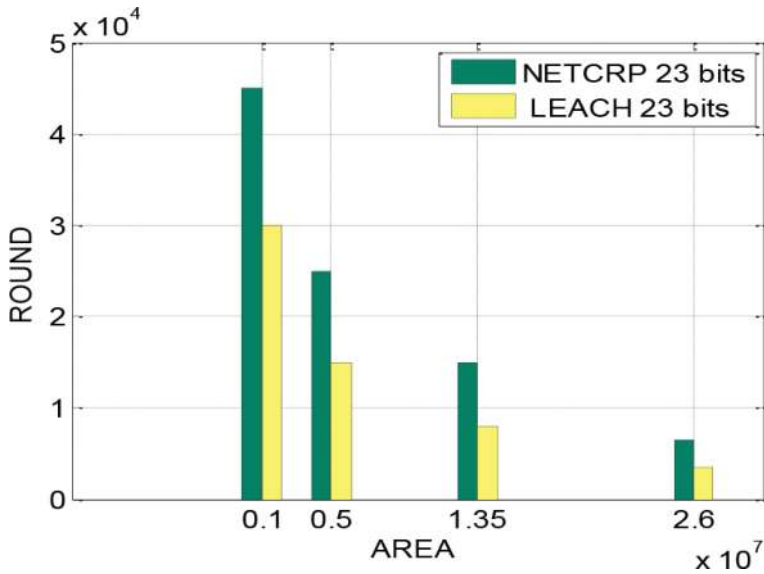


Figure 5. Round in which all the nodes become dead [21].

These observations conclude that the flexible threshold selection method improves WSN network-lifetime by increasing energy savings with respect to earlier methods. Moreover, it is suitable for automated monitoring for all area sizes, large number of nodes and if amount of information to be transmitted is large.

4. Conclusions

This chapter discusses the flexible threshold selection method for efficient environment monitoring of the offshore wind farm. It uses degree of similarity between the previous and the current datasets for calculating geometric mean-based flexible threshold as it does not get biased due to extreme values in the datasets. The method is compared with the static threshold mean method of threshold selection and the performance is seen to be enhanced.

Also, the automated fault detection method is presented in this chapter. This is a simple method that uses small integer values for indicating faulty conditions of the wind farm in real time. The method uses fuzzy inference system that takes integer values as input and gives output in the form of fault status prediction of the farm. Based on these predictions, the system suggests corrective or preventive measures. The method is proved to be very accurate in predicting the fault condition based on sensed parameter values. In addition, this method allows reduction in the size of the transmitted data packets to 23 bits, which help in increasing the overall network lifetime of the WSN system deployed in the wind farm.

Acknowledgements

I would like to thank Amity University, Lucknow, for providing me the means and resources in writing this chapter. I sincerely thank our Pro Vice Chancellor and Director ASET, Lucknow campus..

Author details

Deepshikha Agarwal

Address all correspondence to: sun.mita@gmail.com

Amity University, Lucknow, India

References

- [1] Wisner RH, Bolinger, M. Wind Technologies Market Report. Energy Efficiency & Renewable Energy, U.S. Department of Energy, Springfield, DOE/GO-102014-4459. August 2014. pp. 1-95.
- [2] www.eia.gov/energyexplained. [Accessed on February 1, 2017]
- [3] www.davidsuzuki.org/issues/climate-change. [Accessed on February 1, 2017]
- [4] www.ucsusa.org/clean-energy/coal-and-other-fossil-fuels. [Accessed on February 1, 2017]
- [5] Lau BCP, Ma EWM. Review of offshore wind turbine failures and fault prognostic methods. In: Proceedings of IEEE Conference on Prognostics and System Health Management; Beijing; May 23–25, 2012. pp. 1-5
- [6] Hansen MH. How Hard Can it be to Pitch a Wind Turbine Blade? Silkeborg, Denmark: RISO Lab, Denmark Technical University; 2007
- [7] Teknik FS. Advantages of Hydraulic Pitch Control. Available from: www.fst.dk/dyn/files/normal_items/1029-file-2/paper.Pdf [Accessed: August 15, 2016]
- [8] Mohanty AR, Kar C. Fault detection in a multistage gearbox by demodulation of motor current waveform. IEEE Transactions on Industrial Electronics. June 2006;**53**(4):1285-1297
- [9] Zhang Z, Verma A, Kusiak A. Fault analysis of the wind turbine gearbox. IEEE Transactions on Energy Conversion. 2012;**27**:526-535
- [10] Brandão RF, Carvalho JA, Maciel Barbosa FP. Forecast of faults in a wind turbine gearbox. In: ELEKTRO 2012. Rajec Teplice, Slovakia.
- [11] Badihi H. Model-based active fault-tolerant cooperative control in an offshore wind farm. Energy Procedia. December 2016;**103**:46-51

- [12] An M-S. Implementation of fault diagnosis of wind turbine based on signal analysis with NN algorithm. In: International Conference on Disaster Recovery and Business Continuity (DRBC); Korea; November 25-28, 2015.
- [13] Rajagopalan S, Restrepo JA, Aller J, Habetler T, Harley R. Non stationary motor fault detection using recent quadratic time frequency representations. *IEEE Transaction on Industry Applications*. June 2008;**44**:735-744
- [14] Blodt M, Bonacci D, Regnier J, Chabert M, Faucher J. Online monitoring of mechanical faults in variable-speed induction motor drives using the Wigner distribution. *IEEE Transaction on Industrial Electronics*. February 2008;**55**:522-533
- [15] Riera-Guaspa M, Antonio-Daviu J, Roger-Folch J, Palomares MPM. The use of the wavelet approximation signal as a tool for the diagnosis of rotor bar failure. *IEEE Transaction on Industry Applications*. June 2008;**44**:716-726
- [16] Amirat Y. Wind turbines condition monitoring and fault diagnosis using generator current amplitude demodulation. In: 2010 IEEE International Energy Conference, Bahrain, December 2010
- [17] Al-Ahmar E, Benbouzid MEH, Amirat Y. DFIG-based wind turbine fault diagnosis using a specific discrete wavelet transform. In: International Conference on Electrical Machines, Portugal, September 2008
- [18] Agarwal D, Kishor N. Network lifetime enhanced tri-level clustering and routing protocol for monitoring of off shore wind farms. *IET Wireless Sensor Systems*. 2014;**4**:69-79. DOI: 10.1049/iet-wss.2013.0050
- [19] Agarwal D, Kishor N. Flexible threshold selection and fault prediction method for health monitoring of off-shore wind farm. *IET Wireless Sensor Systems*. 2015;**5**:183-192. DOI: 10.1049/iet-wss.2014.0008
- [20] Agarwal D, Kishor N. A flexible threshold selection and fault detection method for health monitoring of offshore wind farms. In: 2nd International Conference on Information and Communication Technology; Bandung, Indonesia; May 28-30, 2014. 2014. pp. 493-498
- [21] Agarwal D, Kishor N. A fuzzy inference based fault detection scheme using adaptive thresholds for health monitoring of offshore wind farms. *Special Edition on IEEE Sensors Journal*. 2014;**14**:3851-3861. DOI: 10.1109/JSEN.2014.2347700